

10.767.674
07.20.04
日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出 願 年 月 日 2 0 0 3 年 1 1 月 2 7 日
Date of Application:

出 願 番 号 特 願 2 0 0 3 - 3 9 6 7 8 6
Application Number:
[J P 2 0 0 3 - 3 9 6 7 8 6]
ST. 10/C] :

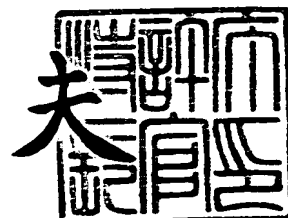
願 人 株式会社日立製作所
Applicant(s):

CERTIFIED COPY OF
PRIORITY DOCUMENT

2 0 0 4 年 1 月 2 7 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



【書類名】 特許願
【整理番号】 340301148
【あて先】 特許庁長官殿
【国際特許分類】 G06F 03/06
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 箕輪 信幸
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 100095371
 【弁理士】
 【氏名又は名称】 上村 輝之
【選任した代理人】
 【識別番号】 100089277
 【弁理士】
 【氏名又は名称】 宮川 長夫
【選任した代理人】
 【識別番号】 100104891
 【弁理士】
 【氏名又は名称】 中村 猛
【手数料の表示】
 【予納台帳番号】 043557
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1
 【包括委任状番号】 0110323

【書類名】 特許請求の範囲**【請求項 1】**

複数のプロセッサを用いて情報を処理する装置において、
1 又は複数の第 1 ローカルメモリを有する 1 以上の第 1 プロセッサと、
ターゲットの第 1 プロセッサが有するターゲット第 1 ローカルメモリにライト情報を直接ライトする、及び／又は、前記ターゲット第 1 ローカルメモリからリード情報を直接リードする 1 以上の第 2 プロセッサと
を備える情報処理装置。

【請求項 2】

前記 1 以上の第 1 プロセッサの各々についての第 1 ローカルメモリアドレスが記録された第 1 アドレスマップを記憶する第 1 アドレスマップ記憶手段を更に備え、

前記 1 以上の第 2 プロセッサの各々が、前記第 1 アドレスマップから前記ターゲット第 1 プロセッサの第 1 ローカルメモリアドレスを取得し、前記取得した第 1 ローカルメモリアドレスに前記ライト情報をライトする、及び／又は、前記取得した第 1 ローカルメモリアドレスから前記リード情報をリードする、
請求項 1 記載の情報処理装置。

【請求項 3】

前記 1 以上の第 2 プロセッサの各々が持つ 1 又は複数の第 2 ローカルメモリと、
前記 1 以上の第 1 プロセッサの各々についての第 1 ローカルメモリアドレスが記録された第 1 アドレスマップを記憶する第 1 アドレスマップ記憶手段と、
前記 1 以上の第 2 プロセッサの各々についての第 2 ローカルメモリアドレスが記録された第 2 アドレスマップを記憶する第 2 アドレスマップ記憶手段と
を更に備え、

前記 1 以上の第 2 プロセッサの各々が、前記ターゲット第 1 ローカルメモリのどこにライトするかの第 1 ローカルメモリライトアドレスを前記第 1 アドレスマップから取得し、前記取得した第 1 ローカルメモリライトアドレスにライト情報をライトし、

前記ターゲット第 1 プロセッサが、前記第 1 ローカルメモリライトアドレスにライトされた前記ライト情報がリードコマンドの場合、前記リードコマンドに応答して、前記リードコマンドの発行元のターゲット第 2 プロセッサの第 2 ローカルメモリライトアドレスを前記第 2 アドレスマップから取得し、前記取得した第 2 ローカルメモリライトアドレスに、前記第 1 ローカルメモリ内の情報を読み出してライトする、
請求項 1 記載の情報処理装置。

【請求項 4】

前記第 2 プロセッサ側から前記ライト情報を受けて前記ターゲット第 1 プロセッサ側に転送する中継デバイスを備え、

前記中継デバイスは、中継メモリを備え、前記ライト情報の転送を行う場合、該ライト情報を前記中継メモリに一時的に格納してから前記転送を行うことと、該ライト情報を前記中継メモリに格納せずに前記転送を行うこととを選択的に実行する、
請求項 1 記載の情報処理装置。

【請求項 5】

前記 1 以上の第 1 プロセッサを有する 1 又は複数の第 1 デバイスを備え、

前記中継デバイスは、前記 1 又は複数の第 1 デバイスに接続され、前記受信したライト情報を、前記 1 又は複数の第 1 デバイスにそれぞれ送信する 1 又は複数の送信部を更に備え、或る第 2 デバイスから受信したライト情報を、前記ターゲット第 1 プロセッサを有するターゲット第 1 デバイスに送信する場合、そのライト情報を送信するターゲット送信部がビジー状態でなければ、前記受信したライト情報を前記中継メモリに格納せずに前記ターゲット送信部から前記ターゲット第 1 デバイスに送信し、前記ターゲット送信部がビジー状態であれば、前記ライト情報を前記中継メモリに一時的に格納し、前記ターゲット送信部のビジー状態が解除された場合に、前記中継メモリからライト情報を読み出して前記ターゲット送信部から前記ターゲット第 1 デバイスに送信する、

請求項 4 記載の情報処理装置。

【請求項 6】

前記中継メモリには、1 以上の送信元又は送信先にそれぞれ対応した 1 以上のライト情報格納領域が用意されており、

前記中継デバイスは、受信した前記ライト情報を前記中継メモリに一時的に格納する場合、そのライト情報を、その送信元又は送信先に対応したターゲットライト情報格納領域に格納し、その場合に、前記ターゲットライト情報格納領域に蓄積された情報量が第 1 の閾値を超えたならば、その旨を表す第 1 閾値超過通知を所定の第 2 デバイスに送信し、

前記第 1 閾値超過通知を受けた第 2 デバイスは、前記ターゲットローカルメモリ又は前記ターゲット第 1 プロセッサに対するライト情報の発行頻度又は情報量を減らす、
請求項 4 記載の情報処理装置。

【請求項 7】

前記中継メモリには、1 以上の送信元又は送信先にそれぞれ対応した 1 以上のライト情報格納領域が用意されており、

前記中継デバイスは、受信した前記ライト情報を前記中継メモリに一時的に格納する場合、該ライト情報を、その送信元又は送信先に対応したターゲットライト情報格納領域に格納し、その場合に、前記ターゲットライト情報格納領域に蓄積された情報量が第 2 の閾値を超えたならば、その旨を表す第 2 閾値超過通知を所定の第 2 デバイスに送信し、

前記第 2 デバイスは、前記ライト情報を前記ターゲットローカルメモリに直接ライトする直接ライト方式と、前記ライト情報を前記中継メモリに格納し前記ターゲット第 1 プロセッサが前記中継メモリから前記ライト情報を取得できるようにする間接ライト方式とを選択的に実行するようになっており、前記第 2 閾値超過通知を受けない場合は、前記直接ライト方式で前記ライト情報を送信し、前記第 2 閾値超過通知を受けた場合は、前記間接ライト方式で前記ライト情報を送信する、

請求項 4 記載の情報処理装置。

【請求項 8】

前記中継デバイスは、前記ターゲットライト情報格納領域に蓄積された情報量が、前記第 1 の閾値よりも大きい第 2 の閾値を超えたならば、その旨を表す第 2 閾値超過通知を前記第 2 デバイスに送信し、

前記第 2 デバイスは、前記ライト情報を前記ターゲットローカルメモリに直接ライトする直接ライト方式と、前記ライト情報を前記中継メモリに格納し前記ターゲット第 1 プロセッサが前記中継メモリから前記ライト情報を取得できるようにする間接ライト方式とを選択的に実行するようになっており、前記第 1 閾値超過通知を受けても前記第 2 閾値超過通知を受けない場合は、前記直接ライト方式で前記ライト情報を送信し、前記第 2 閾値超過通知を受けた場合は、前記直接ライト方式を止めて前記間接ライト方式で前記ライト情報を送信する、

請求項 6 記載の情報処理装置。

【請求項 9】

前記第 2 閾値超過通知を受けた第 2 デバイスが前記間接ライト方式を選択した後、以下の (1) 及び (2) の場合、

(1) 前記ターゲットライト情報格納領域の情報量が前記第 2 の閾値よりも小さい第 3 の閾値以下になった場合、

(2) 前記間接ライト方式が選択されている状態が一定時間経過した場合の少なくとも 1 つの場合に、該第 2 デバイスは前記間接ライト方式を止めて前記直接ライト方式を選択する、

請求項 7 記載の情報処理装置。

【請求項 10】

前記 1 以上の第 1 プロセッサを有する 1 又は複数の第 1 デバイスと、

前記 1 以上の第 2 プロセッサを有する 1 又は複数の第 2 デバイスと、

前記第 2 デバイスから前記ターゲット第 1 プロセッサを有する前記第 1 デバイスに前記

ライト情報を中継する中継デバイスと

を備え、

前記中継デバイスは、

前記 1 又は複数の第 1 デバイスにそれぞれ情報を送信する 1 又は複数の送信部と、

前記 1 又は複数の第 2 デバイスからそれぞれ情報を受信する 1 又は複数の受信部と

を備え、前記 1 又は複数の送信部の各々と、前記 1 又は複数の受信部の各々とが互いに独立して動作する、

請求項 1 記載の情報処理装置。

【請求項 1 1】

前記 1 以上の第 1 プロセッサの各々のローカルメモリには、前記 1 以上の第 2 プロセッサにそれぞれ対応した 1 以上のローカル格納領域が用意されており、

前記 1 以上の第 2 プロセッサの各々は、その第 2 プロセッサに対応した前記ローカル格納領域のローカルメモリアドレスが各第 1 プロセッサ毎に記録されたアドレスマップを記憶しており、前記ライト情報をターゲットの第 1 プロセッサのローカルメモリに書き込む場合、前記アドレスマップから前記ターゲット第 1 プロセッサに対応したローカルメモリアドレスを取得し、前記取得したローカルメモリアドレスに前記ライト情報をライトする、

請求項 1 記載の情報処理装置。

【請求項 1 2】

前記ターゲット第 1 プロセッサのローカルメモリアドレスを有する前記ライト情報を前記第 2 プロセッサ側から受信して前記ターゲット第 1 プロセッサ側に転送する中継デバイスを備え、

前記 1 以上の第 2 プロセッサの各々は、前記中継デバイスを介して、論理的な又は物理的な 1 以上のバスでそれぞれ前記 1 以上の第 1 プロセッサと通信可能に接続されており、

前記中継デバイスは、前記 1 以上のバスにそれぞれ対応付けられた 1 以上のローカルメモリアドレスを各第 2 プロセッサ毎に記憶しており、前記受信したライト情報の転送を行う場合、そのライト情報に含まれている前記ローカルメモリアドレスに対応したターゲットバスを特定し、特定されたターゲットバスを介して前記ライト情報を前記ターゲット第 1 プロセッサ側に転送する、

請求項 1 記載の情報処理装置。

【請求項 1 3】

複数のマイクロプロセッサと物理的又は論理的な記憶装置とを備え、前記複数のマイクロプロセッサを用いて、上位装置からの情報を前記記憶装置に記憶させることの記憶制御を行う記憶制御装置において、

1 又は複数の第 1 ローカルメモリを有する 1 以上の第 1 マイクロプロセッサと、

1 以上の第 2 マイクロプロセッサと、

前記 1 以上の第 1 マイクロプロセッサの各々についての第 1 ローカルメモリアドレスが記録された第 1 アドレスマップを記憶する第 1 アドレスマップ記憶部と

を備え、

前記 1 以上の第 2 マイクロプロセッサの各々が、ターゲット第 1 ローカルメモリのどこにライトするか第 1 ローカルメモリライトアドレスを前記第 1 アドレスマップから取得し、前記取得した第 1 ローカルメモリライトアドレスにライト情報をライトする、

記憶制御装置。

【請求項 1 4】

前記 1 以上の第 1 マイクロプロセッサを搭載した 1 又は複数の第 1 デバイスと、

前記 1 以上の第 2 マイクロプロセッサを搭載した 1 又は複数の第 2 デバイスと、

前記 1 又は複数の第 1 デバイスと前記 1 又は複数の第 2 デバイスと間の通信を中継する中継デバイスと

を備え、

前記第 1 ローカルメモリには、前記 1 以上の第 2 マイクロプロセッサにそれぞれ対応し

た 1 以上のローカル格納領域が用意されており、

前記第 1 アドレスマップには、前記 1 以上の第 1 マイクロプロセッサにそれぞれ対応した 1 以上のローカル格納領域の第 1 ローカルメモリアドレスが記録されており、

前記 1 以上の第 2 マイクロプロセッサの各々は、前記中継デバイスを介して、論理的な又は物理的な 1 以上のバスでそれぞれ前記 1 以上の第 1 マイクロプロセッサと接続可能であり、前記ライト情報を出力する際、前記第 1 アドレスマップから前記ターゲット第 1 プロセッサに対応した第 1 ローカル格納領域の第 1 ローカルメモリアドレスを取得して、前記取得した第 1 ローカルメモリアドレスを含んだ第 1 送信先情報を有するライト情報を出力し、

前記中継デバイスは、前記 1 以上のバスにそれぞれ対応付けられた 1 以上の第 2 送信先情報を記憶しており、前記受信したライト情報の転送を行う場合、前記第 1 と第 2 の送信先情報に基づいてターゲットバスを特定し、特定されたターゲットバスを介して前記ライト情報を前記ターゲット第 1 デバイスに転送し、

前記ターゲット第 1 デバイスは、前記中継デバイスから受信したライト情報を、そのライト情報に含まれている前記第 1 ローカルメモリアドレスに書き込む、
請求項 13 記載の記憶制御装置。

【請求項 15】

前記中継デバイスは、中継メモリを備え、前記ライト情報の転送を行う場合、そのライト情報を前記中継メモリに一時的に格納してから前記転送を行うことと、そのライト情報を前記中継メモリに格納せずに前記転送を行うこととを選択的に実行する、
請求項 14 記載の記憶制御装置。

【請求項 16】

前記中継デバイスは、前記 1 又は複数の第 1 デバイスに通信可能に接続され、前記受信したライト情報を前記 1 又は複数の第 1 デバイスにそれぞれ送信する 1 又は複数の送信部を更に備え、或る第 2 デバイスから受信したライト情報を、前記ターゲット第 1 デバイスに送信する場合、前記ターゲットバスがビジー状態でなければ、前記受信したライト情報を前記中継メモリに格納せずに前記ターゲットバスを介して前記ターゲット第 1 デバイスに送信し、前記ターゲットバスがビジー状態であれば、前記ライト情報を前記中継メモリに一時的に格納し、前記ターゲットバスのビジー状態が解除された場合に、前記中継メモリからライト情報を読み出して前記ターゲットバスを介して前記ターゲット第 1 デバイスに送信する、

請求項 14 記載の記憶制御装置。

【請求項 17】

前記中継メモリには、ライト情報の 1 以上の送信元又は送信先にそれぞれ対応した 1 以上のライト情報格納領域が用意されており、

前記中継デバイスは、受信した前記ライト情報を前記中継メモリに一時的に格納する場合、そのライト情報を、その送信元又は送信先に対応したターゲットライト情報格納領域に格納し、その場合に、前記ターゲットライト情報格納領域に蓄積された情報量が第 1 の閾値を超えたならば、その旨を表す第 1 閾値超過通知を所定の第 2 デバイスに送信し、

前記第 1 閾値超過通知を受けた第 2 デバイスは、前記ターゲットローカルメモリ又は前記ターゲット第 1 マイクロプロセッサに対するライト情報の発行頻度又は情報量を減らす、
請求項 15 記載の記憶制御装置。

【請求項 18】

前記中継デバイスは、

前記 1 又は複数の第 1 デバイスにそれぞれ情報を送信する 1 又は複数の送信部と、

前記 1 又は複数の第 2 デバイスからそれぞれ情報を受信する 1 又は複数の受信部とを備え、前記 1 又は複数の送信部の各々と、前記 1 又は複数の受信部の各々とが互いに独立して動作する、

請求項 14 記載の記憶制御装置。

【請求項 19】

前記中継デバイスは、前記中継メモリとは別の受信バッファを備えており、

前記 1 又は複数の受信部の各々は、前記第 2 デバイスから受信したライト情報を一旦前記受信バッファに格納する、
請求項 18 記載の記憶制御装置。

【請求項 20】

複数のマイクロプロセッサと物理的又は論理的な記憶装置とを備え、前記複数のマイクロプロセッサを用いて、上位装置からの情報を前記記憶装置に記憶させることの記憶制御を行う記憶制御装置において、

1 又は複数の第 1 ローカルメモリを有する 1 以上の第 1 マイクロプロセッサと、

1 又は複数の第 2 ローカルメモリを有する 1 以上の第 2 マイクロプロセッサと、

前記 1 以上の第 1 マイクロプロセッサの各々についての第 1 ローカルメモリアドレスが記録された第 1 アドレスマップを記憶する第 1 アドレスマップ記憶手段と、

前記 1 以上の第 2 マイクロプロセッサの各々についての第 2 ローカルメモリアドレスが記録された第 2 アドレスマップを記憶する第 2 アドレスマップ記憶手段とを備え、

前記 1 以上の第 2 マイクロプロセッサの各々が、ターゲット第 1 ローカルメモリのどこにライトするか第 1 ローカルメモリライトアドレスを前記第 1 アドレスマップから取得し、前記取得した第 1 ローカルメモリライトアドレスにリードコマンドをライトし、

前記ターゲット第 1 マイクロプロセッサが、前記第 1 ローカルメモリライトアドレスにライトされた前記リードコマンドに应答して、前記リードコマンドの発行元のターゲット第 2 マイクロプロセッサの第 2 ローカルメモリライトアドレスを前記第 2 アドレスマップから取得し、前記取得した第 2 ローカルメモリライトアドレスに、前記第 1 ローカルメモリ内のリード情報を読み出してライトする、
記憶制御装置。

【請求項 21】

複数のプロセッサを用いて情報を処理する方法において、

1 以上の第 2 プロセッサの各々が、1 又は複数のローカルメモリを有する 1 以上の第 1 プロセッサの各々についてのローカルメモリアドレスが記録されたアドレスマップから、ターゲット第 1 プロセッサのローカルメモリアドレスを取得するステップと、

前記 1 以上の第 2 プロセッサの各々が、前記取得したローカルメモリアドレスにライト情報をライトする、及び／又は、前記取得したローカルメモリアドレスからリード情報をリードするステップとを有する情報処理方法。

【書類名】明細書**【発明の名称】**複数のプロセッサを用いて情報処理を行う装置及び方法**【技術分野】****【0001】**

本発明は、複数のプロセッサを用いて情報処理を行うための技術に関し、具体的には、例えば、RAID (Redundant Array of Independent Inexpensive Disks) システムに備えられる記憶制御装置に関する。

【背景技術】**【0002】**

複数のプロセッサを用いて情報処理を行う装置の一つに、例えば、アレイ状に配列された複数の物理ディスクを有する記憶装置に対するホスト装置からのアクセスを制御する記憶制御装置がある。記憶制御装置には、例えば、ホスト装置と通信するための2以上のチャネルアダプタと、記憶装置と通信するための2以上のディスクアダプタとが備えられており、各アダプタには、1又は複数のマイクロプロセッサ（以下、「MP」と略記する）が備えられている。

【0003】

チャネルアダプタに搭載されているMP（以下、「チャネルMP」と言う）と、ディスクアダプタに搭載されているMP（以下、「ディスクMP」と言う）の間では制御情報がやり取りされる。例えば特許文献1に開示の技術によれば、チャネルMP17とディスクMP21との間での制御情報のやり取りは、共有メモリ32を介して行われる。具体的には、例えば、チャネルMP17がディスクMP21に制御情報を送る場合、チャネルMP17が、共有メモリ32の所定の領域に制御情報をライトし、一方、ディスクMP21が、共有メモリ32の所定領域を定期的にポーリングし、その所定領域に制御情報があれば、その所定領域から制御情報を取得する。

【0004】

【特許文献1】 特開2001-306265（例えば段落44）

【発明の開示】**【発明が解決しようとする課題】****【0005】**

上述した従来技術によれば、MP間での制御情報のやり取りは、共有メモリを介して行われており、受信側のMPが、共有メモリの所定領域を定期的にポーリングする。共有メモリをポーリングすることは、受信側MPにとって少なからず負担となる。また、共有メモリを複数MPが頻繁にリードを繰り返すことは、MP数が多くなるほど、共有メモリの性能ボトルネック、共有メモリバスの性能ボトルネックが発生しやすく、それ故、装置全体の性能を低下させてしまう可能性がある。

【0006】

従って、本発明の目的は、受信側プロセッサにとって負担が少ない方法で、プロセッサ間での情報の授受が行えるようにすることにある。

【0007】

本発明の別の目的は、装置全体の性能を低下させることなく、プロセッサ間での情報の授受が行えるようにすることにある。

【0008】

本発明の他の目的は、後述の説明から明らかになるであろう。

【課題を解決するための手段】**【0009】**

本発明の第1の側面に従う情報処理装置は、1又は複数のローカルメモリを有する1以上の第1プロセッサと、ターゲットの第1プロセッサが有するローカルメモリにライト情報を直接ライトする、及び／又は、前記ターゲット第1ローカルメモリからリード情報（リード対象の情報）を直接リードする1以上の第2プロセッサとを備える。

【0010】

ここで、「プロセッサ」とは、例えば、演算処理を行うモジュール、具体例としては、CPU (Central Processing Unit) 或いはMP (Micro Processor) がある。

【0011】

また、「ローカルメモリ」とは、第1プロセッサのみが使用可能なメモリであって、例えば、CPUが利用するRAM等のメモリ、或いは、MPの内部又は外部に搭載されたメモリ (RAM或いはそれとは別のキャッシュメモリ) 等がある。

【0012】

また、「ライト情報」は、例えば、ターゲット第1プロセッサに対するメッセージ、又は、その第1プロセッサのローカルメモリから情報をリードすることを要求するリードコマンドである。

【0013】

この発明の第1の好適な実施形態では、前記1以上の第1プロセッサの各々についてのローカルメモリアドレスが記録されたアドレスマップを記憶するアドレスマップ記憶手段を更に備え、前記1以上の第2プロセッサの各々が、前記アドレスマップから前記ターゲット第1プロセッサのローカルメモリアドレスを取得し、前記取得したローカルメモリアドレスに前記ライト情報をライトする、及び/又は、前記取得した第1ローカルメモリアドレスから前記リード情報をリードする。

【0014】

この発明の第2の好適な実施形態では、前記1以上の第2プロセッサの各々が持つ1又は複数の第2ローカルメモリと、前記1以上の第1プロセッサの各々についての第1ローカルメモリアドレスが記録された第1アドレスマップを記憶する第1アドレスマップ記憶手段 (これは、例えば、第2プロセッサが有しても良い) と、前記1以上の第2プロセッサの各々についての第2ローカルメモリアドレスが記録された第2アドレスマップを記憶する第2アドレスマップ記憶手段 (これは、例えば、第1プロセッサが有しても良い) とが更に備えられる。前記1以上の第2プロセッサの各々が、前記ターゲット第1ローカルメモリのどこにライトするか第1ローカルメモリライトアドレスを取得し (例えば、更に、前記ターゲット第1ローカルメモリのどこからリードするか第1ローカルメモリリードアドレスを前記第1アドレスマップから取得し)、前記取得した第1ローカルメモリライトアドレスに所定のリードコマンド (例えば、前記取得した第1ローカルメモリリードアドレスを指定するリードコマンド) をライトする。前記ターゲット第1プロセッサが、前記第1ローカルメモリライトアドレスにライトされた前記リードコマンドに応答して、前記リードコマンドの発行元のターゲット第2プロセッサの第2ローカルメモリライトアドレスを前記第2アドレスマップから取得し、前記取得した第2ローカルメモリライトアドレスに、前記第1ローカルメモリ内の情報 (例えば、前記リードコマンドから識別される前記第1ローカルメモリリードアドレス内の情報) をライトする。

【0015】

この発明の第3の好適な実施形態では、情報処理装置が、第2プロセッサ側から前記ライト情報を受けて前記ターゲット第1プロセッサ側に転送する中継デバイスを備える。前記中継デバイスは、中継メモリを備え、前記ライト情報の転送を行う場合、該ライト情報を前記中継メモリに一時的に格納してから前記転送を行うことと、該ライト情報を前記中継メモリに格納せずに前記転送を行うこととを選択的に実行する。

【0016】

この発明の第4の好適な実施形態では、情報処理装置が、前記1以上の第1プロセッサを有する1又は複数の第1デバイスを備える。前記中継デバイスは、前記1又は複数の第1デバイスに通信可能に接続され、前記受信したライト情報を前記1又は複数の第1デバイスにそれぞれ送信する1又は複数の送信部を更に備え、或る第2デバイスから受信したライト情報を、前記ターゲット第1プロセッサを有するターゲット第1デバイスに送信する場合、そのライト情報を送信するターゲット送信部がビジー状態でなければ、前記受信したライト情報を前記中継メモリに格納せずに前記ターゲット送信部から前記ターゲット第1デバイスに送信し、前記ターゲット送信部がビジー状態であれば、前記ライト情報を

前記中継メモリに一時的に格納し、前記ターゲット送信部のビジー状態が解除された場合に、前記中継メモリからライト情報を読み出して前記ターゲット送信部から前記ターゲット第1デバイスに送信する。

【0017】

ここで、「送信部」とは、受信したライト情報を前記ターゲット第1デバイスに送信することができるものならどのようなものでも良く、例えば、通信ポート或いはバスが考えられる。

【0018】

この発明の第5の好適な実施形態では、前記第3の好適な実施形態において、前記中継メモリには、ライト情報の1以上の送信元又は送信先にそれぞれ対応した1以上のライト情報格納領域が用意されており、前記中継デバイスは、受信した前記ライト情報を前記中継メモリに一時的に格納する場合、そのライト情報を、その送信元又は送信先に対応したターゲットライト情報格納領域に格納し、その場合に、前記ターゲットライト情報格納領域に蓄積された情報量が第1の閾値を超えたならば、その旨を表す第1閾値超過通知を所定の第2デバイスに送信する。前記第1閾値超過通知を受けた第2デバイスは、前記ターゲットローカルメモリ又は前記ターゲット第1プロセッサに対するライト情報の発行頻度又は情報量を減らす。

【0019】

ここで、「送信元」とは、例えば、ライト情報を出力する第2プロセッサ又はそれを搭載した第2デバイスであり、「送信先」とは、例えば、ライト情報の格納先となるローカルメモリ又はそれを有する第1プロセッサである。

【0020】

「ライト情報の発行頻度又は情報量を減らす」方法として、例えば、以下の2つの方法がある。

【0021】

(1) 前記中継デバイスが、前記ライト情報送信元の第2プロセッサ、又はそれを含む全ての第2プロセッサに対して所定の第1割り込み信号を送信し、その第1割り込み信号を受けた第2プロセッサそれぞれ自身が、前記ターゲット第1プロセッサに対するライト情報の発行頻度又は情報量を減らす。

【0022】

(2) 前記第1閾値超過通知を受けた第2プロセッサそれぞれ自身は、前記ターゲット第1プロセッサに対するライト情報の発行頻度又は情報量を減らさず、前記1以上の第2プロセッサを搭載した第2デバイスにおける情報送信制御部が、第2プロセッサから出力されたライト情報を受け、そのライト情報が、前記ターゲット第1プロセッサのローカルメモリを宛先とするものであれば、そのライト情報を前記中継デバイスに送信することを中止或いは中断し、一方、そのライト情報が、前記ターゲット第1プロセッサのローカルメモリを宛先とするものでなければ、そのライト情報を前記中継デバイスに送信する。

【0023】

この発明の第6の好適な実施形態では、前記第3の好適な実施形態において、前記中継メモリには、前記1以上の送信元又は送信先にそれぞれ対応した1以上のライト情報格納領域が用意されており、

前記中継デバイスは、受信した前記ライト情報を前記中継メモリに一時的に格納する場合、該ライト情報を、その送信元又は送信先に対応したターゲットライト情報格納領域に格納し、その場合に、前記ターゲットライト情報格納領域に蓄積された情報量が第2の閾値を超えたならば、その旨を表す第2閾値超過通知を所定の第2デバイスに送信する。前記第2デバイスは、前記ライト情報を前記ターゲットローカルメモリに直接ライトする直接ライト方式と、前記ライト情報を前記中継メモリに格納し前記ターゲット第1プロセッサが前記中継メモリから前記ライト情報を取得できるようにする間接ライト方式とを選択的に実行するようになっており、前記第2閾値超過通知を受けない場合は、前記直接ライト方式で前記ライト情報を送信し（例えば、前記ターゲット第1プロセッサのローカル

メモリアドレスを含んだライト情報を送信し)、前記第2閾値超過通知を受けた場合は、前記間接ライト方式で前記ライト情を送信する(例えば、前記中継メモリにおける前記ターゲット第1格納領域に対応した中継メモリアドレスを含んだライト情報を送信する)。

【0024】

この発明の第7の好適な実施形態では、前記第5の好適な実施形態において、前記中継デバイスは、前記ターゲットライト情報格納領域に蓄積された情報量が、前記第1の閾値よりも大きい第2の閾値を超えたならば、その旨を表す第2閾値超過通知を前記第2デバイスに送信する。前記第2デバイスは、前記ライト情報を前記ターゲットローカルメモリに直接ライトする直接ライト方式と、前記ライト情報を前記中継メモリに格納し前記ターゲット第1プロセッサが前記中継メモリから前記ライト情報を取得できるようにする間接ライト方式とを選択的に実行するようになっており、前記第1閾値超過通知を受けても前記第2閾値超過通知を受けない場合は、前記直接ライト方式で前記ライト情報を送信し、前記第2閾値超過通知を受けた場合は、前記直接ライト方式を止めて前記間接ライト方式で前記ライト情報を送信する。

【0025】

この発明の第8の好適な実施形態では、前記第6の好適な実施形態において、前記第2閾値超過通知を受けた第2デバイスが前記間接ライト方式を選択した後、以下の(1)及び(2)の場合、

(1) 前記ターゲットライト情報格納領域の情報量が前記第2の閾値よりも小さい第3の閾値以下になった場合、

(2) 前記間接ライト方式が選択されている状態が一定時間経過した場合の少なくとも1つの場合に、該第2デバイスは前記間接ライト方式を止めて前記直接ライト方式を選択する。

【0026】

この発明の第9の好適な実施形態では、情報処理装置は、前記1以上の第1プロセッサを有する1又は複数の第1デバイスと、前記1以上の第2プロセッサを有する1又は複数の第2デバイスと、前記第2デバイスから前記ターゲット第1プロセッサを有する前記第1デバイスに前記ライト情報を中継する中継デバイスとを備える。前記中継デバイスは、前記1又は複数の第1デバイスにそれぞれ情報を送信する1又は複数の送信部と、前記1又は複数の第2デバイスからそれぞれ情報を受信する1又は複数の受信部とを備える。前記1又は複数の送信部の各々と、前記1又は複数の受信部の各々とは、互いに独立して動作する。例えば、1又は複数の送信部と1又は複数の受信部との間に、1以上の記憶デバイス(例えば揮発性又は不揮発性メモリ)が用意され、1又は複数の送信部と1又は複数の受信部との間でやり取りされる情報が前記1以上の記憶デバイスに格納されることにより、上記互いに独立した動作が可能となる。

【0027】

この発明の第10の好適な実施形態では、前記1以上の第1プロセッサの各々のローカルメモリには、前記1以上の第2プロセッサにそれぞれ対応した1以上のローカル格納領域が用意されている。前記1以上の第2プロセッサの各々は、その第2プロセッサに対応した前記ローカル格納領域のローカルメモリアドレスが各第1プロセッサ毎に記録されたアドレスマップを記憶しており、前記ライト情報をターゲットの第1プロセッサのローカルメモリに書き込む場合、前記アドレスマップから前記ターゲット第1プロセッサに対応したローカルメモリアドレスを取得し、前記取得したローカルメモリアドレスに前記ライト情報をライトする。

【0028】

この発明の第11の好適な実施形態では、情報処理装置は、前記ターゲット第1プロセッサのローカルメモリアドレスを含んだ前記ライト情報を前記第2プロセッサ側から受信して前記ターゲット第1プロセッサ側に転送する中継デバイスを備える。前記1以上の第2プロセッサの各々は、前記中継デバイスを介して、論理的な又は物理的な1以上のパスでそれぞれ前記1以上の第1プロセッサと通信可能に接続されている。前記中継デバイス

は、前記1以上のパスにそれぞれ対応付けられた1以上のローカルメモリアドレスを各第2プロセッサ毎に記憶しており、前記受信したライト情報の転送を行う場合、そのライト情報に含まれている前記ローカルメモリアドレスに対応したターゲットパスを特定し、特定されたターゲットパスを介して前記ライト情報を前記ターゲット第1プロセッサ側に転送する。

【0029】

本発明の第2の側面に従う記憶制御装置は、複数のマイクロプロセッサと物理的又は論理的な記憶装置とを備え、前記複数のマイクロプロセッサを用いて、上位装置からの情報を前記記憶装置に記憶させることの記憶制御を行う記憶制御装置であって、1又は複数の第1ローカルメモリを有する1以上の第1マイクロプロセッサと、1以上の第2マイクロプロセッサと、前記1以上の第1マイクロプロセッサの各々についての第1ローカルメモリアドレスが記録された第1アドレスマップを記憶する第1アドレスマップ記憶手段とを備える。前記1以上の第2マイクロプロセッサの各々が、前記ターゲット第1ローカルメモリのどこにライトするか第1ローカルメモリライトアドレスを取得し、前記取得した第1ローカルメモリライトアドレスにライト情報をライトする。

【0030】

この発明の第1の好適な実施形態では、前記1以上の第1マイクロプロセッサを搭載した1又は複数の第1デバイスと、前記1以上の第2マイクロプロセッサを搭載した1又は複数の第2デバイスと、前記1又は複数の第1デバイスと前記1又は複数の第2デバイスと間の通信を中継する中継デバイスとが更に備えられる。前記第1ローカルメモリには、前記1以上の第2マイクロプロセッサにそれぞれ対応した1以上のローカル格納領域が用意されている。前記第1アドレスマップには、前記1以上の第1マイクロプロセッサにそれぞれ対応した1以上のローカル格納領域の第1ローカルメモリアドレスが記録されている。前記1以上の第2マイクロプロセッサの各々は、前記中継デバイスを介して、論理的な又は物理的な1以上のパスでそれぞれ前記1以上の第1マイクロプロセッサと通信可能に接続されていて、前記ライト情報を出力する際、前記第1アドレスマップから前記ターゲット第1プロセッサに対応した第1ローカル格納領域の第1ローカルメモリアドレスを取得して、前記取得した第1ローカルメモリアドレスを含んだ第1送信先情報（例えば、ローカルメモリアドレスのみ、或いは更にライト情報の宛先マイクロプロセッサ識別情報を含んだ情報）を有するライト情報を出力する。前記中継デバイスは、前記1以上のパスにそれぞれ対応付けられた1以上の第2送信先情報（例えばローカルメモリアドレス又はそれに対応付けられた宛先マイクロプロセッサ識別情報）を記憶しており、前記受信したライト情報の転送を行う場合、前記第1と第2の送信先情報に基づいてターゲットパスを特定し、特定されたターゲットパスを介して前記ライト情報を前記ターゲット第1デバイスに転送する。前記ターゲット第1デバイスは、前記中継デバイスから受信したライト情報を、そのライト情報に含まれている前記第1ローカルメモリアドレスに書き込む。

【0031】

この発明の第2の好適な実施形態では、前記中継デバイスは、中継メモリを備え、前記ライト情報の転送を行う場合、そのライト情報を前記中継メモリに一時的に格納してから前記転送を行うことと、そのライト情報を前記中継メモリに格納せずに前記転送を行うこととを選択的に実行する。

【0032】

この発明の第3の好適な実施形態では、第2の好適な実施形態において、前記中継デバイスは、前記1又は複数の第1デバイスに通信可能に接続され、前記受信したライト情報を前記1又は複数の第1デバイスにそれぞれ送信する1又は複数の送信部を更に備え、或る第2デバイスから受信したライト情報を、前記ターゲット第1デバイスに送信する場合、前記ターゲットパスがビジー状態でなければ、前記受信したライト情報を前記中継メモリに格納せずに前記ターゲットパスを介して前記ターゲット第1デバイスに送信し、前記ターゲットパスがビジー状態であれば、前記ライト情報を前記中継メモリに一時的に格納し、前記ターゲットパスのビジー状態が解除された場合に、前記中継メモリからライト情

報を読み出して前記ターゲットパスを介して前記ターゲット第1デバイスに送信する。

【0033】

この発明の第4の好適な実施形態では、第2の好適な実施形態において、前記中継メモリには、ライト情報の1以上の送信元又は送信先にそれぞれ対応した1以上のライト情報格納領域が用意されている。前記中継デバイスは、受信した前記ライト情報を前記中継メモリに一時的に格納する場合、そのライト情報を、その送信元又は送信先に対応したターゲットライト情報格納領域に格納し、その場合に、前記ターゲットライト情報格納領域に蓄積された情報量が第1の閾値を超えたならば、その旨を表す第1閾値超過通知を所定の第2デバイスに送信する。前記第1閾値超過通知を受けた第2デバイスは、前記ターゲットローカルメモリ又は前記ターゲット第1マイクロプロセッサに対するライト情報の発行頻度又は情報量を減らす。

【0034】

ここで、「送信元」とは、例えば、ライト情報を出力する第2マイクロプロセッサ又はそれを搭載した第2デバイスであり、「送信先」とは、例えば、ライト情報の格納先となるローカルメモリ又はそれを有する第1マイクロプロセッサである。

【0035】

この発明の第5の好適な実施形態では、第2の好適な実施形態において、前記中継デバイスは、前記1又は複数の第1デバイスにそれぞれ情報を送信する1又は複数の送信部と、前記1又は複数の第2デバイスからそれぞれ情報を受信する1又は複数の受信部とを備え、前記1又は複数の送信部の各々と、前記1又は複数の受信部の各々が互いに独立して動作する。

【0036】

この発明の第6の好適な実施形態では、第5の好適な実施形態において、前記中継デバイスは、前記中継メモリとは別の受信バッファを備えており、前記1又は複数の受信部の各々は、前記第2デバイスから受信したライト情報を一旦前記受信バッファに格納する。

【0037】

本発明の第3の側面に従う記憶制御装置は、複数のマイクロプロセッサと物理的又は論理的な記憶装置とを備え、前記複数のマイクロプロセッサを用いて、上位装置からの情報を前記記憶装置に記憶させることの記憶制御を行う記憶制御装置であって、1又は複数の第1ローカルメモリを有する1以上の第1マイクロプロセッサと、1又は複数の第2ローカルメモリを有する1以上の第2マイクロプロセッサと、前記1以上の第1マイクロプロセッサの各々についての第1ローカルメモリアドレスが記録された第1アドレスマップを記憶する第1アドレスマップ記憶手段と、前記1以上の第2マイクロプロセッサの各々についての第2ローカルメモリアドレスが記録された第2アドレスマップを記憶する第2アドレスマップ記憶手段とを備える。前記1以上の第2マイクロプロセッサの各々が、ターゲット第1ローカルメモリのどこにライトするか第1ローカルメモリライトアドレスを前記第1アドレスマップから取得し、前記取得した第1ローカルメモリライトアドレスにリードコマンドをライトする。前記ターゲット第1マイクロプロセッサが、前記第1ローカルメモリライトアドレスにライトされた前記リードコマンドに応答して、前記リードコマンドの発行元のターゲット第2マイクロプロセッサの第2ローカルメモリライトアドレスを前記第2アドレスマップから取得し、前記取得した第2ローカルメモリライトアドレスに、前記第1ローカルメモリ内のリード情報を読み出してライトする。

【0038】

なお、上述した「ライト情報」又は「リード情報」は、例えば、プロセッサ間でやり取りされるデータ固定長（又はデータ可変長の）メッセージ（換言すれば制御情報）であっても良いし、ユーザに指定された書き込み対象又は読み出し対象のユーザ所望情報（例えば、アプリケーションソフト等によって作成されたファイル）であっても良い。

【発明の効果】

【0039】

本発明に従う情報処理装置及び記憶制御装置によれば、第2プロセッサが、ターゲット

の第1プロセッサが有するローカルメモリにライト情報を直接ライトする、及び／又は、そのローカルメモリからリード情報を直接リードする。例えば、直接ライトの場合、ターゲット第1プロセッサは、自分が持つローカルメモリをポーリングすることで、第2プロセッサから出力されたライト情報を取得することができる。第1プロセッサがローカルメモリをリードする時間長は、共有メモリをリードする時間長に比べて短い（例えば10分の1以下程度である）。このため、第1プロセッサは、負担少なく、第2プロセッサから出力されたライト情報を取得することができる。換言すれば、従来のような共有メモリの定期リードによるプロセッサ、共有メモリ及び共有メモリパスの性能の低下を防止することができる。

【発明を実施するための最良の形態】

【0040】

以下、図面を参照して、本発明の一実施形態に係る情報処理装置が適用されたディスクアレイ制御装置について説明する。

【実施例1】

【0041】

図1は、本発明の第一実施例に係るディスクアレイ制御装置の構成を示す。

【0042】

この図に示すディスクアレイ制御装置100は、例えば、RAID (Redundant Array of Independent Inexpensive Disks) 装置であり、この装置100には、1又は複数のホストコンピュータ200と、複数の物理ディスク群9とが接続される。

【0043】

ホストコンピュータ（以下、適宜「ホスト」と略記）200は、例えば、パーソナルコンピュータやワークステーション等であり、CPU (Central Processing Unit) やメモリ等を備えたコンピュータマシンである。ホスト200のCPUが各種プログラムを実行することにより、種々の機能が実現される。ホスト200は、例えば、LAN等の通信ネットワークを介して、ディスクアレイ制御装置100に対し、所望のデータを物理ディスク群9（例えば論理デバイス又は論理ユニットと呼ばれる論理記録領域）にライトするライトコマンドや、所望のデータを物理ディスク群9からリードするリードコマンドをディスクアレイ制御装置100に送信する。

【0044】

物理ディスク群9は、アレイ状に配設された複数のディスク型記憶装置600を有する。物理ディスク群9上には、物理的な記憶領域又は論理的な記憶領域（例えば論理デバイス又は論理ユニットと呼ばれる論理記録領域）が設けられ、各記憶領域に、ホストコンピュータ200からのデータがライトされる。なお、ディスク型記憶装置600としては、例えば、ハードディスク装置、フレキシブルディスク装置、半導体記憶装置等の種々のデバイスを用いることができる。

【0045】

ディスクアレイ制御装置100は、ホストコンピュータ200の複数の物理ディスク群9に対するアクセスを制御するための装置である。具体的には、ディスクアレイ制御装置100は、ホストコンピュータ200からリード要求を受けて、リード要求されたデータを物理ディスク群9（例えば所定の論理記憶領域）から取得してそれをホストコンピュータ200に転送したり、ホストコンピュータ200からライトデータと共にライト要求を受けて、そのライトデータを物理ディスク群9に格納したりする。ディスクアレイ制御装置100には、複数（又は1つ）のホストアダプタ300と、複数（又は1つ）のディスクアダプタ500と、複数のホストアダプタ300と複数のディスクアダプタ500とに接続される中継アダプタ400とが備えられる。

【0046】

各ホストアダプタ300には、1又は複数台のホストコンピュータ200とLAN等の通信ネットワークを介して通信可能に接続され、且つ、中継アダプタ400に所定の通信インターフェースを介して通信可能に接続される（例えばスター接続される）。各ホスト

アダプタ 300 は、ハードウェア回路（又は、ソフトウェア或いはそれらの組み合わせ）で構成することができ、このディスクアレイ制御装置 100 とそれに接続されているホストコンピュータ 200 との間の通信を制御する。各ホストアダプタ 300 は、接続されているホスト 200 と通信を行うための通信インターフェースを有し、また、ホスト 200 から受信した各種コマンドを解釈して処理するためのコマンドプロセッサ機能も有する。具体的には、各ホストアダプタ 300 は、1 又は複数のマイクロプロセッサ（以下、「ホスト側 MP」と言う）360 と、各ホスト側 MP 360 に専用の 1 以上（例えば 1 つの）ローカルメモリ（以下、「ホスト側 LM」と言う）350 と、各ホスト側 MP 360 と中継アダプタ 400 とを通信可能に接続するためのホストハブ（換言すれば、例えば、ホスト側 MP 360 の中継アダプタ 400 に対するアダプタ）310 とを備える。

【0047】

各ディスクアダプタ 500 には、1 台（又は複数台）の物理ディスク群 9 と所定の通信インターフェース（例えば SAN (Storage Area Network) 等の通信ネットワーク）を介して通信可能に接続され、且つ、中継アダプタ 400 に所定の通信インターフェースを介して通信可能に接続される（例えばスター接続される）。各ディスクアダプタ 500 は、ハードウェア回路（又は、ソフトウェア或いはそれらの組み合わせ）で構成することができ、このディスクアレイ制御装置 100 とそれに接続されている物理ディスク群 9 との間の通信を制御する。具体的には、各ディスクアダプタ 500 は、1 又は複数のマイクロプロセッサ（以下、「ディスク側 MP」と言う）550 と、各ディスク側 MP 550 に専用の 1 以上のローカルメモリ（以下、「ディスク側 LM」と言う）560 と、各ディスク側 MP 550 と中継アダプタ 400 とを通信可能に接続するためのディスクハブ（換言すれば、例えば、ディスク側 MP 550 の中継アダプタ 400 に対するアダプタ）510 とを備える。

【0048】

中継アダプタ 400 は、複数のアダプタ 300、500 間の通信を制御するためのポートコントローラ 410 と、複数の MP 360、550 間で共有可能な共有メモリ（以下、Shared Memory を略して「SM」と言う）470 とを備える。SM 470 には、直接 LM ライト方式のときにメッセージが一時的に格納されるバッファ領域と、間接 LM ライト方式のときにメッセージが格納されるキャッシュ領域とが含まれている。なお、バッファ領域及びキャッシュ領域のうち少なくともバッファ領域には、各 MP 360 及び 550 毎に対応したサブバッファ領域が用意されている。このため、例えば、或るホスト側 MP 360 が直接 LM ライト方式で出力したメッセージが一時的に SM 470 に格納される場合、メッセージ送信元のホスト側 MP 360 に対応したサブバッファ領域にメッセージが格納される。なお、上述した「直接 LM ライト方式」とは、メッセージの送信側 MP 360（又は 550）が、受信側 MP 550（又は 360）の介在無しに、その MP 550（又は 360）が有する LM 560（350）にメッセージをライトする方式を言う。一方、「間接 LM ライト方式」とは、実質的には共有メモリ方式のことであり、送信側 MP 360（又は 550）が SM 470 のキャッシュ領域にメッセージをライトし、受信側 MP 550（又は 360）がそのキャッシュ領域をポーリングすることにより、その受信側 MP 550（又は 360）がメッセージを取得し自分の LM 560（350）にライトする方式を言う。

【0049】

以上のような構成の下で、ホスト側 MP 360 が、ディスク側 MP 550 の介在無しで、中継アダプタ 400 のポートコントローラ 410 を経由して、所定の又は所望のディスク側 LM 560 にメッセージ（MP 間でやり取りされる制御情報）を直接書き込むことができる。同様に、ディスク側 MP 550 が、ホスト側 MP 360 の介在無しで、中継アダプタ 400 のポートコントローラ 410 を経由して、所望のホスト側 LM 350 にメッセージを直接書き込むことができる。これを可能とした一つの工夫として、各 MP 360、550 に、少なくとも 1 つの他の MP 360 又は 550 が有する少なくとも 1 つの LM 350 又は 560 のアドレス（以下、「LM アドレス」と言う）が記録されたアドレスマッ

プが格納されていることが挙げられる。それにより、各MP 360、550は、そのアドレスマップから、所望の又は所定のLMアドレスを取得し、そのLMアドレスに、中継アダプタ400のポートコントローラ410を介してメッセージを書き込むことができる。以下、この点について具体的に説明する。

【0050】

図2は、各MP 360（及び550）に格納されているアドレスマップの一例を示す。

【0051】

各MP 360（及び550）、例えば、MP識別情報（「MP#」と図示）が「00」であるホストMP 360に格納されているアドレスマップ4には、種々の情報が記録されているが、主要な情報の例として、送信時SMアドレス情報、受信時SMアドレス情報、他MPのLMアドレス情報、及び自MPのLMアドレス情報がある。

【0052】

送信時SMアドレス情報は、ホスト側MP 360が間接LMライト方式で他のMP 550（又は360）にメッセージを送信する場合に、SM 470のどのアドレスにメッセージを格納すべきかを示す情報である。送信時SMアドレス情報は、例えば、所定の1又は複数の他MP 550（又は360）の識別情報と、それら1又は複数の他MPにそれぞれ対応した1又は複数のSMアドレス範囲（1又は複数のSMアドレス）を含んでいる。

【0053】

受信時SMアドレス情報は、他MP 550（又は360）から間接LMライト方式でメッセージを受信する場合に、SM 470のどのアドレスをリードすべきかを示す情報である。受信時SMアドレス情報は、例えば、所定の1又は複数の他MP 550（又は360）の識別情報と、それら1又は複数の他MPにそれぞれ対応した1又は複数のSMアドレス範囲（1又は複数のSMアドレス）を含んでいる。

【0054】

他MPのLMアドレス情報は、所定の1又は複数の他MPがそれぞれ有するLMのアドレスを含んだ情報（図示のように、自MPのLMアドレスが含まれていても良い）である。具体的には、例えば、他MPのLMアドレス情報には、図2に示すように、1以上（例えば8個）のアダプタ300及び500の識別情報（例えばアダプタ名）と、1以上のアダプタ300、500に搭載されている1以上（例えばMP#「00」を除いた15個）の他MP 550（又は360）の識別情報と、それら1以上の他MP 550（又は360）が有する1以上のLM 560（又は350）のLMアドレス範囲（直接ライトすることが可能な1又は複数のLMアドレス）とが含まれている。

【0055】

自MPのLMアドレス情報は、自分が有する1以上のLM 350（又は560）のLMアドレス範囲（1又は複数のLMアドレス）を含んだ情報である。具体的には、例えば、図示のように、MP#「00」のホスト側MP 360のメッセージの宛先となるディスク側MP 550（例えばMP#「08」）の自MPのLMアドレス情報には、1以上の送信元MP 360又は550にそれぞれ対応したLMアドレス範囲（1又は複数のLMアドレス）が含まれている。

【0056】

以上のように、各MP 360、550のプログラムアドレス空間に、上記のようなアドレスマップ4を記録しておく。これにより、そのMP 360（又は550）のマイクロプログラムの処理によって、任意のターゲットMP 550（又は360）のLMアドレスにライト命令を実行すると、ターゲットMP 550（又は360）のLM 560（又は350）にメッセージが書き込まれる。すなわち、MP 360（又は550）は、宛先MP 550（又は360）の介在無しで、中継アダプタ400のポートコントローラ410を経由して、所定の又は所望のディスク側LM 560にメッセージを直接書き込むことができる。具体的には、例えば、MP#「00」のホスト側MP 360がMP#「08」のディスク側MP 550にメッセージを発行する場合、そのホスト側MP 360は、自分が有するアドレスマップ4から、MP#「08」のディスク側MP 550に対応したLMアドレ

スを取得し、そのLMアドレスと所望メッセージとを含んだ後述の直接ライトコマンドを生成して、中継アダプタ400に送信する。すると、その直接ライトコマンドは、中継アダプタ400から、MP#「08」のディスク側MP550を有するディスクアダプタ500に転送される。そして、その直接ライトコマンドに含まれるメッセージは、MP#「08」のディスク側MP550のディスク側LM560の、送信元ホスト側MP360（MP#「00」のMP）に対応したLMアドレスに格納される。

【0057】

なお、図示のベースアドレスは、LM560（又は350）のオフセットを示したもので、ターゲットMP550（又は360）が搭載されているアダプタ500（又は300）のハブ510（又は310）の所与の設定値で、あらかじめターゲットMP550（又は360）がマイクロプログラムにより設定しておくものである。

【0058】

以下、ホスト側MP360が所望のディスク側LM560に対するメッセージを発行する場合を例に、ホストアダプタ300、中継アダプタ400及びディスクアダプタ500の構成及び機能について詳細に説明する。

【0059】

図3は、ホストアダプタ300の機能及び構成を詳細に示すブロック図である。図4は、中継アダプタ400の機能及び構成を詳細に示すブロック図である。図5は、ディスクアダプタ500の機能及び構成を詳細に示すブロック図である。図3、4及び5を横に並べて見ることで、アダプタ300、400及び500の繋がりが分かるようになっている。

【0060】

図3に示すように、ホストアダプタ300には、1以上（例えば2つ）のホスト側MP360A、360Bと、それら1以上のホスト側MP360A、360Bにそれぞれ対応した1以上（例えば2つ）のホスト側LM350A、350Bが搭載されている。ホストアダプタ300では、所定のプロセッサバス（例えばPCI-Xバス）2を介して、1以上のホスト側MP360A、360B、ホスト側LM350A、350B、及びホストハブ310が通信可能に接続されている。ホストハブ310には、MPバス制御部311、アクセス制御部316、1以上のホスト側MP360A、360Bにそれぞれ対応した1以上（例えば2つ）のホストポート200A、200B、ステータス判定部317、SMバッファ閾値テーブル314、SMバッファフルテーブル315、割込み制御部313及びリセット制御部312が備えられている。

【0061】

MPバス制御部311は、LMアドレス及びメッセージを含むコマンド（以下、「直接ライトコマンド」と言う）やSMアドレス及びメッセージを含むコマンド（以下、「間接ライトコマンド」と言う）をホスト側MP360A、360Bから受信し、そのコマンドをアクセス制御部316に送信する。また、MPバス制御部311は、後述するリセット命令をホスト側MP360から受信した場合、そのリセット命令をリセット制御部312に通知する。

【0062】

アクセス制御部316は、MPバス制御部311から受信したライトコマンドを、その送信元のホスト側MP360A又は360Bに対応したホストポート200A又は200Bに転送する。その際、アクセス制御部316は、SMバッファ閾値テーブル314及びSMバッファフルテーブル315に書かれている情報に基づいて、MPバス制御部311から受けたコマンドをホストポート200A又は200Bに送るか否かを制御する。具体的には、例えば、アクセス制御部316は、SMバッファフルテーブル315にバッファフルが記録されていないホスト側MP360Aからの直接ライトコマンドはホストポート360Aに送信するが、バッファフルが記録されているホスト側MP360Bからの直接ライトコマンドはホストポート360Bに送信せずに止める。

【0063】

各ホストポート 200A、200B には、送信制御部 318 と受信制御部 319 とが含まれている。

【0064】

送信制御部 318 は、アクセス制御部 316 から受信したコマンドを所定のフォーマットに変換して中継アダプタ 400 のポートコントローラ 410 に送信する。フォーマット変換後のコマンドは、図 6 に示す通りである。すなわち、送信制御部 318 からポートコントローラ 410 に送られるコマンドは、アドレス情報、コマンド種別、ライトデータ及び CRC を含んでいる。アドレス情報は、例えば、LM アドレス又は SM アドレスである。コマンド種別には、例えば、共有メモリにメッセージをライトすることを意味する「共有メモリライト」、共有メモリからメッセージをリードすることを意味する「共有メモリリード」、及び、メッセージを直接 LM にライトすることを意味する「LM ライト」がある（更に、メッセージのデータ長などが含まれていても良い）。ライトデータは、送信元 MP 360 又は 550 で生成されたメッセージである。CRC は、このコマンドフォーマットのデータ保証コードを示す。

【0065】

受信制御部 319 は、ポートコントローラ 410 からステータス情報を受信し、ステータス判定部 317 に渡す。ステータス情報は、例えば図 7 に示す通りである。すなわち、ステータス情報には、種々の情報が含まれているが、主要な情報の例として、宛先 MP 情報、バッファフル情報及びバッファ閾値オーバー情報がある。宛先 MP 情報は、例えば、このステータス情報が送信される起因となったメッセージの宛先の MP 550 の識別情報である。バッファフル情報は、宛先 MP 550 に対応した SM 470 上のサブバッファ領域で後述するバッファフル（換言すれば、例えば重度の警告）が発生したか否かを表す情報である。バッファ閾値オーバー情報は、宛先 MP 550 に対応した SM 470 上でのサブバッファ領域で後述するバッファ閾値オーバー（換言すれば、例えば軽度の警告）が発生したか否かを表す情報である。

【0066】

SM バッファ閾値テーブル 314 には、全ての MP 360、550 の各々について、それに対応するサブバッファ領域に存在する直接ライトコマンドの蓄積量が第 1 閾値を超えたか否か（換言すれば、バッファ閾値オーバーが発生したか否か）が記録される。具体的には、例えば、或る宛先 MP 550 に対応したサブバッファ領域で第 1 閾値を超えていない場合には、SM バッファ閾値テーブル 314 における、その MP 550 に対応したビットはオフになっており、そのサブバッファ領域で第 1 閾値を超えた場合には、ステータス判定部 317 によって、その MP 550 に対応したビットはオンにされる。このとき、その旨が SM バッファ閾値テーブル 314 から割り込み制御部 313 及びアクセス制御部 316 に通知される。

【0067】

また、SM バッファ閾値テーブル 314 には、閾値登録代表ビットが備えられる。閾値登録代表ビットは、全ての MP 360 及び 550 にそれぞれ対応した全てのビットがオフであれば、オフであるが、少なくとも 1 つの MP 360 又は 550 に対応したビットがオンになればオンになる。

【0068】

SM バッファフルテーブル 315 には、全ての MP 360、550 の各々について、それに対応するサブバッファ領域に存在する直接ライトコマンドの蓄積量が第 1 閾値よりも大きい第 2 閾値を超えたか否か（換言すれば、バッファフルが発生したか否か）が記録される。具体的には、例えば、或る宛先 MP 550 に対応したサブバッファ領域で第 2 閾値を超えていない場合には、SM バッファフルテーブル 315 における、その MP 550 に対応したビットはオフになっており、そのサブバッファ領域で第 2 閾値を超えた場合には、ステータス判定部 317 によって、その MP 550 に対応したビットはオンにされる。このとき、その旨が SM バッファフルテーブル 315 から割り込み制御部 313 及びアクセス制御部 316 に通知される。

【0069】

SMバッファフルテーブル315には、閾値登録代表ビットが備えられる。閾値登録代表ビットは、全てのMP360及び550にそれぞれ対応した全てのビットがオフであれば、オフであるが、少なくとも1つのMP360又は550に対応したビットがオンになればオンになる。

【0070】

ステータス制御部317は、受信制御部319から受けたステータス情報（図7参照）の宛先MP情報、バッファフル情報及びバッファ閾値オーバー情報に基づいて、SMバッファ閾値テーブル314及びSMバッファフルテーブル315を更新する。具体的には、例えば、ステータス判定部317は、バッファフル情報がバッファフルを示していれば、宛先MP情報が示すMP550についてのSMバッファフルテーブル315上のビットをオフからオフにする。同様に、例えば、ステータス判定部317は、バッファ閾値オーバー情報が閾値オーバーを示していれば、宛先MP情報が示すMP550についてのSMバッファ閾値テーブル314上のビットをオフからオフにする。

【0071】

割込み制御部313は、2種類のテーブル314及び315にそれぞれ記録されている各MP360、550毎の情報に基づいて、各MP360に割り込み信号を出力するか否か、出力するのであればどのような割り込み信号を出力するかを制御する。具体的には、例えば、割込み制御部313は、或るディスク側MP550について閾値オーバーが発生した旨の通知をSMバッファ閾値テーブル314から受けた場合、この割込み制御部313が搭載されているホストアダプタ300上にあるメッセージ送信元MP（又は、全てのMP）360A及び／又は360Bに正常割込み信号を送信する。また、例えば、割込み制御部313は、或るディスク側MP550についてバッファフルが発生した旨の通知をSMバッファフルテーブル315から受けた場合、この割込み制御部313が搭載されているホストアダプタ300上にあるメッセージ送信元MP（又は、全てのMP）360A及び／又は360Bに障害割込み信号を送信する。なお、正常割込み信号を受けたMP360A又は360Bのマイクロプログラムは、上記ディスク側MP550についての閾値オーバーをリセットするための閾値オーバーリセット命令をリセット制御部312に発行する。また、障害割込み信号を受けたMP360のマイクロプログラムは、所定のイベントが発生するまで（例えば、バッファフルが解除された旨の通知をポートコントローラ410から受けるまで）、上記ディスク側MP550を宛先としたメッセージの発行頻度を減らすか、或いは、直接LMライト方式から間接LMライト方式に切り替えて上記ディスク側MP550を宛先とするメッセージを発行し、所定のイベントが発生したら（例えば間接LMライト方式を実行してから一定時間が経過したら）、通常通りに直接LMライト方式でその宛先にメッセージを発行する。

【0072】

リセット制御部312は、ホスト側MP360からMPバス制御部311を介して閾値オーバーリセット命令を受け、命令されたMP550（又は360）についてのSMバッファ閾値テーブル314のビットをオンからオフにする（すなわち、そのMP550についての閾値オーバー状態を解除する）。また、リセット制御部312は、後述する方法によって、或るMP550（又は360）についてのバッファフルリセット命令を受け、そのMP550（又は360）についてのSMバッファフルテーブル315のビットをオンからオフにする（すなわち、そのMP550（又は360）についてのバッファフル状態を解除する）。

【0073】

さて、次に、図4を参照して中継アダプタ400について説明する。

【0074】

中継アダプタ400には、前述したように、ポートコントローラ410とSM470とが備えられる。ポートコントローラ410には、1以上のホストアダプタ300がそれぞれ接続される1以上のホスト側ポートセット420と、1以上のディスクアダプタ500

がそれぞれ接続される1以上のディスク側ポートセット430と、ポート選択制御部440とが備えられる。また、ポートコントローラ410には、リモート直接LMライト制御部450と、メモリ制御部460とが備えられる。

【0075】

ホスト側ポートセット420には、各ホストアダプタ300の1以上のポート200A、200Bにそれぞれ接続される1以上（例えば2つ）のホスト側ポート420A、420Bが備えられる。ホスト側ポート420A、420Bは、ホストアダプタ300と通信するための通信インターフェースである。ホスト側ポート420A、420Bには、接続先ホストポート200A、200Bの送信制御部318からフォーマット変換されたコマンドを受信する受信制御部421と、後述する方法で作成されたステータス情報を接続先ホストポート200A、200Bの受信制御部319に送信する送信制御部422とが備えられる。

【0076】

ディスク側ポートセット430には、各ディスクアダプタ500の1以上のディスクポート520A、520Bにそれぞれ接続される1以上（例えば2つ）のディスク側ポート430A、430Bが備えられる。ディスク側ポート430A、430Bは、ディスクアダプタ500と通信するための通信インターフェースである。ディスク側ポート430A、430Bには、ホスト側ポート420A又は420Bからポート選択制御部440を介して受信したコマンドを接続先ディスクポート520A、520Bの受信制御部に送信する等を行う送信制御部432と、接続先ディスクポート520A、520Bの送信制御部からコマンドを受信する受信制御部とが備えられる。その送信制御部が出力するコマンドのフォーマットは、例えば図8に示すようになっている。すなわち、そのコマンドは、アドレス情報（例えばLMアドレス）と、コマンド種別と、宛先MPのLMにライトすべきライトデータ（すなわち宛先MPへのメッセージ）とが含まれている。なお、例えば、送信制御部432は、送信するコマンドにLMアドレスを含める際、そのLMアドレスの上位Nビットを取り除いても良い。その場合、例えば、上位Nビットが除かれたLMアドレスは、ディスクハブ510において、LM560のベースアドレスを使用して正確なLMアドレスに復元される。

【0077】

ポート選択制御部440は、例えば、ホストアダプタ300から直接ライトコマンドを受信したとき、そのコマンドに含まれているLMアドレスから、そのコマンドをどのディスク側ポート430を介して送信すれば良いのかを識別する。そして、その際、ポート選択制御部440は、識別されたディスク側ポート430が使用中か否か（具体的には、例えば、ディスク側ポート430Aが別のコマンドを送信中である又は受信であるか否か）を判断し、使用中でなければ、直接ライトコマンドをそのポート430Aを介してディスクアダプタ500に送信し、使用中であれば、直接ライトコマンドをSM470上のサブバッファ領域に一時保存するために、直接ライトコマンドのLMアドレスをアドレスデコーダ451に通知し、且つ、直接ライトコマンド内のLMアドレス及びライトデータをデータとしてメモリ制御部460に送信する。また、ポート選択制御部440は、使用中であったポート430Aが未使用になった場合に、ポート430Aからリード要求を受け、リード要求元のポート430Aを特定することで、リードすべきメッセージの宛先MPの識別情報（例えば宛先MPの番号）を識別し、その識別情報をアドレスデコーダ451に通知する。

【0078】

図9は、ポート選択制御部440の概要を示す。この図を参照して、ポート選択制御部440を詳細に説明する。

【0079】

ポート選択制御部440は、1以上のバッファ601A～601nと、ポート管理テーブル600と、パス切替制御部650とを備える。

【0080】

1以上のバッファ601A~601nは、1以上のホスト側ポート420A~420n（又は1以上のディスク側ポート430A~430n）にそれぞれ対応したバッファである。各バッファ601A~601nには、そのバッファに対応したホスト側ポート（又はディスク側ポート）が受信した直接ライトコマンドが一時的に格納される。具体的には、例えば、バッファ601Aには、ホスト側ポート420Aがホストアダプタ300から受信した直接ライトコマンド、及び、或るディスク側ポートからホスト側ポート420Aに転送される直接ライトコマンドが格納される。

【0081】

ポート管理テーブル600には、1以上のホスト側ポート420A~420n及び1以上のディスク側ポート430A~430nの各々毎に、そのポートが受信した直接ライトコマンドをどのポートに転送するのかに関するポート制御情報が登録されている。ポート制御情報は、例えば、1以上のポート430A~430n及び420A~420nの識別情報にそれぞれ対応した1以上の格納先LMアドレス範囲（及びMP識別情報）を含んでいる。

【0082】

パス切替制御部650は、各ポート420A~420n及び430A~430nから受信した直接ライトコマンドを別のポートから出力するためのハードウェア回路である（ソフトウェア又はそれらの組合せであっても良い）。パス切替制御部650は、例えば以下の（1）~（3）の処理、

（1）各ポート420A~420n及び430A~430nから受信された直接ライトコマンドを、そのポートに対応したバッファ601A~601nに格納する処理、

（2）そのバッファに格納された直接ライトコマンドを読み出し、そのコマンドに含まれているLMアドレスに対応したポートから出力する処理、

（3）出力先ポートが使用中であれば、直接ライトコマンドをSM470上のサブバッファ領域に一時保存するために、直接ライトコマンドのLMアドレスをアドレスデコーダ451に通知し、且つ、直接ライトコマンド内のLMアドレス及びライトデータをデータとしてメモリ制御部460に送信する処理、

を行えるのであればどのような構成であっても良い。具体例として、例えば、図10に示すような構成を採用しても良い。すなわち、パス切替制御部650は、1以上のディスク側ポート430A~430nにそれぞれ対応した1以上のポートパス制御部602A~602nと、SMパス制御部603とを備えても良い。1以上のポートパス制御部602A~602nの各々は、1以上のバッファ601A~601nと、自分に対応したディスク側ポートとの間の通信を制御するハードウェア機器（例えばセクタ）である。SMパス制御部603は、1以上のバッファ601A~601nと、アドレスデコーダ451及びメモリ制御部460との間の通信を制御するハードウェア機器（例えばセクタ）である。

。

【0083】

図10に示した構成では、例えば以下の処理が行われる。

【0084】

ポート管理テーブル600には、各ディスク側ポート430A~430nに対応した各ポートパス制御部602A~602nの識別情報や、SMパス制御部603の識別情報が記録されており、それ故、各ホスト側ポート420A~420nが、このテーブル600を参照することにより、受信した直接ライトコマンド中のLMアドレスに基づきどのポートパス制御部602A~602nとアービトレイトすれば良いかを識別したり、SMパス制御部603を識別してそれとアービトレイトしたりすることができるようになっている。

。

【0085】

ホスト側ポート420Bは、受信した直接ライトコマンドをバッファ601Bに格納し、且つ、そのコマンド中のLMアドレスと、ポート管理テーブル600とに基づいて、アービトレイト先となるポートパス制御部602Aを識別し、そのポートパス制御部602

Aとアービトレイトする。

【0086】

アービトレーションの結果、ポートパス制御部602Aが使用中でなければ（例えば、その制御部602Aから直接ライトコマンドを出力することができる状態であれば）、ポートパス制御部602Aが、バッファ601Bに蓄積された直接ライトコマンドを読み出し、それをディスク側ポート430Aに送信する。

【0087】

一方、上記アービトレーションの結果、ポートパス制御部602Aが使用中であれば、ホスト側ポート420Bは、SMパス制御部603とアービトレイトする。その結果、SMパス制御部603が、バッファ601Bに蓄積された直接ライトコマンドを読み出し、そのコマンド内のLMアドレス及びライトデータをデータとしてメモリ制御部460へ転送し、且つ、そのコマンド内のLMアドレスをアドレスデコーダ451に通知する。

【0088】

以上が、ポート制御部440についての説明である。上述の通り、ポート制御部440には、ホスト側ポート420A～420nとディスク側ポート430A～430nとの間にバッファが介在し、各ポート420A～420n及び430A～430nが受信した直接ライトコマンドは一旦バッファに蓄積される。それにより、直接ライトコマンドを受信した各ポート420A～420n及び430A～430nは、そのコマンドを所定のバッファに格納してしまいさえすれば、そのコマンドが所定の出力先ポートから出力されなくても、次に別のコマンドを受けることができる。換言すれば、ホスト側ポート420A～420nとディスク側ポート430A～430nは独立して動作することができる。

【0089】

再び図4を参照する。

【0090】

リモート直接LMライト制御部450は、直接ライトコマンドをSM470上のサブバッファ領域に格納することを制御するものである。リモート直接LMライト制御部450は、スタックポインタテーブル453と、アドレスデコーダ451と、アドレス制御部456と、閾値判定部455と、バッファフル判定部454と、ステータス生成部452とを備える。

【0091】

スタックポインタテーブル453には、メッセージの宛先となる各MP550（及び／又は360）毎にスタックポインタが登録されている。そのスタックポインタは、SM470上の上記バッファ領域（直接ライトコマンドが格納される領域）における、1以上のMPにそれぞれ対応した1以上のサブバッファ領域を指し示すものであって、具体的には、例えば、SM470のベースアドレスからのオフセットを示す。各宛先MP550のスタックポインタには、コマンドのライト位置を示すライトポイントと、コマンドのリード位置を示すリードポイントとがある。スタックポインタテーブル453は、アドレスデコーダ451から或る宛先MP550についてライト要求を受けたときは、そのMP550のライトポイントを1インクリメントし、そのMP550についてリード要求を受けたときは、リードポイント及びライトポイントを1デクリメントし、更新後のライトポイント又はリードポイントと、ポイントの更新対象となったMP550の識別情報とを、アドレス制御部456、バッファフル判定部454及び閾値判定部455に送信する。この実施例では、各サブバッファ領域においてデータが溢れそうか否かは、ライトポイントの値から判別することができる。

【0092】

アドレスデコーダ451は、1以上の格納先LMアドレス範囲にそれぞれ対応した1以上のMP識別情報が記録された図示しないテーブルを記憶している。アドレスデコーダ451は、ポート選択制御部440からLMアドレスが通知された場合、そのLMアドレスと上記テーブルからメッセージの宛先MP550（又は360）を識別し、識別されたMP550に対応するスタックポインタを選択し、スタックポインタテーブル453に対し

、そのMP 550についてのライト要求を送信する。また、アドレスデコーダ451は、ポート選択制御部440から宛先MPの識別情報を受けたときは、その識別情報を持つMP 550に対応するスタックポインタを選択し、スタックポインタテーブル453に対し、そのMP 550についてのリード要求を送信する。

【0093】

アドレス制御部456は、SM 470上のサブバッファ領域に対するSMアドレスを設定する機能を持つ。具体的には、例えば、アドレス制御部456は、スタックポインタテーブル453から受けるライトポインタ又はリードポインタの値と、ポートコントローラ410内に予め設定されているSM 470のベースアドレスとに基づいて、SM 470のどこにデータを格納すべきかを示すSMライトアドレス、又は、SM 470のどこからデータをリードすべきかを示すSMリードアドレスを算出する。アドレス制御部456は、算出されたSMライトアドレス又はSMリードアドレスをメモリ制御部460に通知する。

【0094】

閾値判定部455は、スタックポインタテーブル453から通知されたスタックポインタ（ライトポインタ）が第1閾値を超えたか否かを判定し、超えていれば、閾値オーバーが発生した旨と、それが発生した宛先MP 550の識別情報とをステータス生成部452に送信する。

【0095】

バッファフル判定部454は、スタックポインタテーブル453から通知されたスタックポインタ（ライトポインタ）が第1閾値よりも大きい第2閾値を超えたか否かを判定し、超えていれば、バッファフルが発生した旨と、それが発生した宛先MP 550の識別情報とをステータス生成部452に送信する。なお、その第2閾値は、SM 470上のサブバッファ領域でのデータの蓄積量がそのバッファ領域の記憶容量を超えないようにするための閾値であって、そのバッファ領域の容量以下でその容量に近い値となっており、第1閾値は、その値よりも所定値だけ小さい値となっている。

【0096】

ステータス生成部452は、閾値判定部455から閾値オーバー発生の通知及びそれが発生した宛先MPの識別情報を受けた場合、その宛先MP識別情報と閾値オーバーの発生を示すコードとを含んだステータス情報を生成する。また、ステータス生成部452は、バッファフル判定部454からバッファフル発生の通知及びそれが発生した宛先MPの識別情報を受けた場合、その宛先MP識別情報とバッファフルの発生を示すコードとを含んだステータス情報を生成する。ステータス生成部452は、生成したステータス情報を、ポート選択制御部440及び所定のポート420（及び／又は430）を介して、所定のMP 360、550（例えば、全てのMP 360、550又はメッセージの送信元のホスト側MP 360）に送信する。具体的には、例えば、ステータス生成部452がステータス情報をポート選択制御部440に送信した場合、ポート選択制御部440は、そのステータス情報を、そのステータス情報が生成される起因となった直接ライトコマンドを受信したホスト側ポートから出力する。それにより、そのステータス情報は、その直接ライトコマンドを出力したホストポートが受信し、ホストアダプタ300では、そのホストポートから、どのホスト側MP 360A、360Bに対して正常割込み或いは障害割込みを発行すれば良いかが識別される。

【0097】

メモリ制御部460は、アドレス制御部456からSMライトアドレスを受けた場合、そのSMライトアドレスに、ポート選択制御部440から受けたデータ、すなわち、LMアドレス及びメッセージを格納する。これにより、そのメッセージの宛先MP 550に対応したSM 470上のサブバッファ領域に、そのメッセージ及びそのLMアドレスが格納される。メモリ制御部460は、アドレス制御部456からSMリードアドレスを受けた場合、そのSMリードアドレスにあるデータ（換言すれば、サブバッファ領域に格納されているLMアドレス及びメッセージのセット）を読み出す。

【0098】

さて、次に、図5を参照してディスクアダプタ500について詳細に説明する。

【0099】

図5に示すように、ディスクアダプタ500では、所定のプロセッサバス（例えばPCI-Xバス）6を介して、1以上（例えば2つ）のディスク側MP550A、550B、1以上のディスク側LM560A、560B、及びディスクハブ510が通信可能に接続されている。ディスクハブ510には、中継アダプタ400のディスク側ポート430A、430にそれぞれ接続されるディスクポート520A、520Bと、LMベースアドレス設定部513と、LMアドレス制御部514と、MPバス制御部515とが備えられる。各ディスクポート520A、520Bには、受信制御部516と送信制御部512とが備えられている。

【0100】

受信制御部516は、ポートコントローラ410のディスク側ポートの送信制御部432から、図8に示したフォーマットのコマンド（すなわち、LMアドレス及びメッセージを含んだコマンド）を受信し、それをLMアドレス制御部514に送る。

【0101】

LMベースアドレス設定部513は、各LM560のベースアドレスを記憶している記憶装置（例えば不揮発性のメモリ）である。

【0102】

LMアドレス制御部514は、受信制御部516から受けたLMアドレスと、LMベースアドレス設定部513が記憶しているベースアドレスとを用いて正確なLMアドレスを算出し、そのLMアドレスと、受信制御部516から受けたメッセージとをMPバス制御部515に渡す。

【0103】

MPバス制御部515は、プロセッサバス6を介して、LMアドレス制御部514から受けたLMアドレスに、そのLMアドレス制御部514から受けたメッセージを書き込む。

【0104】

送信制御部512は、ディスク側MP550Aから出力されたコマンドを受けて中継アダプタ400に出力する。

【0105】

以上が、ホストアダプタ300、中継アダプタ400及びディスクアダプタ500についての説明である。

【0106】

なお、上述の説明では、ホスト側MP360がディスク側MP510に対するメッセージを発行する場合を例にとっているが、ディスク側MP510がホスト側MP360に対するメッセージを発行する場合もあるので、ホストハブ310とディスクハブ510は、上述したホストハブ310の機能及び構成と、上述したディスクハブ510の機能及び構成との両方を備えている。すなわち、例えば、ホストハブ310は、ディスクハブ510と同様のLMアドレス制御部514等を備えているし、ディスクハブ510は、ホストハブ310と同様の割込み制御部313等を備えている。

【0107】

また、上述の説明では、ホスト側MP360がディスク側MP510に対するメッセージを発行する場合を例にとっているが、ホスト側MP360は、自分以外のホスト側MP360にメッセージを発行しても良いし、同様に、ディスク側MP550は、自分以外のディスク側MP550にメッセージを発行してもよい。なお、その場合、メッセージは、中継アダプタ400のポートコントローラ410を介してやり取りされても良い。

【0108】

以下、上述したホストアダプタ300、中継アダプタ400及びディスクアダプタ500の間で行われるデータ処理の流れを説明する。

【0109】

本実施例では、各MP 360、550は、他のMP 550、360に対するメッセージを発行する場合、原則として、直接LMライト方式のメッセージを送信する（換言すれば、LMアドレス、コマンド種別「LMライト」及びメッセージを有するコマンドを発行する）。以下、メッセージの送信元MPをホスト側MP 360Aとし、メッセージの宛先MPをディスク側MP 550Aとして、アダプタ300、400及び500が行うデータのやり取りの流れを説明する。

【0110】

(1) 第1ケース：直接LMライト方式のメッセージが送信される場合であって、そのメッセージがSM 470のサブバッファ領域に一時格納されることが無い場合。

【0111】

この第1ケースにおいてアダプタ300、400及び500が行うデータのやり取りの流れを図11に示す。以下、この第1ケースについての流れについて、図11、図3～図5及び図9を参照して説明する。

【0112】

ホスト側MP 360Aは、自分が記憶するアドレスマップ4から、メッセージの格納先LMアドレスであって、宛先MPに対応したLMアドレスを取得し、取得したLMアドレスと、そのLMアドレスにライトすべきメッセージとを含んだライト命令をMPバス2に出力する（ステップS1）。すると、そのライト命令は、MPバス制御部311及びアクセス制御部316を介して、自MP 360Aに対応したホストポート200Aの送信制御部318に送信される。送信制御部318は、そのコマンドを所定のフォーマットに変換し、フォーマット変換後のコマンド（すなわち、LMアドレス、コマンド種別「直接LMライト」及びメッセージを含んだコマンド）を、ホストポート200Aの接続先のホスト側ポート420Aが有する受信制御部421に送信する（S2）。

【0113】

ポートコントローラ410のホスト側ポート420Aが有する受信制御部421は、ホストポート200Aの送信制御部318から受信したコマンドを、ポート選択制御部440内の所定バッファ601Aに一時格納する。また、その受信制御部421は、そのコマンドに含まれているLMアドレスとポート管理テーブル600とに基づいて、どのディスク側ポート430A～430nを使用すべきかを識別し、識別されたターゲットポート430A（つまり、格納先LMアドレスに対応したディスク側ポート）に対して、所定のリクエスト信号を送信する（S3）。ホスト側ポート420Aが、そのリクエスト信号に応答して、ターゲットポート430A（又はそれに対応したポートバス制御部602A）から使用中で無い旨を表す信号（「not busy」と図示）を受けた場合（S4）、すなわち、ターゲットポート430Aが未使用である場合、バッファ601A内のコマンドが読み出されてターゲットポート430Aの送信制御部432を介してディスクアダプタ500に転送される（S5、S6）。また、ホスト側ポート420Aの送信制御部442は、コマンドの転送が完了したことを表す通知を、接続先ホストポート200Aの受信制御部319に送信する（それにより、ホストアダプタ300において、出力したコマンドの転送が終了したことを認識することができる）（S7）。

【0114】

中継アダプタ400からコマンドを受けたディスクポート520Aの受信制御部516は、そのコマンド（すなわち、LMアドレス及びメッセージを含んだコマンド）をLMアドレス制御部514に送る。

【0115】

LMアドレス制御部514は、受信制御部516から受けたLMアドレスと、LMベースアドレス設定部513が記憶しているベースアドレスとを用いて正確なLMアドレスを算出し、そのLMアドレスと、受信制御部516から受けたメッセージとをMPバス制御部515に渡す。

【0116】

MPバス制御部515は、プロセッサバス6を介して、LMアドレス制御部514から受けたLMアドレスに、そのLMアドレス制御部514から受けたメッセージを書き込む(S8)。

【0117】

ディスク側MP550は、自分が記憶しているアドレスマップの自MPのLMアドレス情報に基づいて、自分が有するLM560のうち送信元MP360に対応したLMアドレスをポーリングする(すなわち、そのLMアドレスを定期的にリードする)(S9、S10)。その場合、そのLMアドレスにメッセージがあれば、ディスク側MP550は、そのLMアドレスからメッセージを取得する(S11、S12)。

【0118】

以上の一連の流れにより、ホスト側MP360が発行したメッセージがディスク側LM560に直接ライトされ、直接ライトされたメッセージがディスク側MP550に伝えられる。

【0119】

(2) 第2ケース：直接LMライト方式のメッセージが送信される場合であって、そのメッセージがSM470のサブバッファ領域に一時格納される場合。

【0120】

この第2ケースにおいてアダプタ300、400及び500が行うデータのやり取りの流れを図12に示す。以下、この第2ケースについての流れについて、図12、図3～図5及び図9を参照して説明する。

【0121】

ホスト側ポート420Aの受信制御部421は、上記第1ケースと同様の流れで直接ライトコマンドを受信して(S21、S22)、所定バッファ601Aに一時格納し、そのコマンドに含まれているLMアドレスとポート管理テーブル600とに基づいてターゲットポート430Aを識別し、そのターゲットポート430Aに対して所定のリクエスト信号を送信する(S23)。ホスト側ポート420Aは、そのリクエスト信号に応答して、ターゲットポート430Aから使用中である旨を表す信号(「busy」と図示)を受けた場合(S24)、すなわち、ターゲットポート430が使用中である場合、バス切替制御部650が、受信したコマンド中のLMアドレスをアドレスデコーダ451に通知し、且つ、受信したコマンドに含まれているメッセージ及びLMアドレスをデータ α としてメモリ制御部460に送信する(S25)。その後、ホスト側ポート420Aは、コマンドの転送が完了したことを表す通知を、ホストポート200Aの受信制御部319に送信する(S26)。

【0122】

アドレスデコーダ451は、ポート選択制御部440からLMアドレスが通知された場合、そのLMアドレスからメッセージの宛先MP550を識別して、識別されたMP550に対応するスタックポインタを選択し、スタックポインタテーブル453に対し、そのMP550についてのライト要求を送信する。

【0123】

スタックポインタテーブル453は、アドレスデコーダ451から或る宛先MP550についてライト要求を受けたときは、そのMP550のライトポインタを1インクリメントし、更新後のライトポインタの値と、ポインタの更新対象となったMP550の識別情報とを、アドレス制御部456、バッファフル判定部454及び閾値判定部455に送信する。

【0124】

アドレス制御部456は、スタックポインタテーブル453から受けるライトポインタの値と、ポートコントローラ410内に予め設定されているSM470のベースアドレスとに基づいて、SM470のどこにデータを格納すべきかを示すSMライトアドレスを算出する。アドレス制御部456は、算出されたSMライトアドレスをメモリ制御部460に通知する。

【0125】

メモリ制御部460は、アドレス制御部456からSMライトアドレスを受けた場合、そのSMライトアドレスに、ポート選択制御部440から受けたデータ α 、すなわち、LMアドレス及びメッセージを格納する。

【0126】

その後、ターゲットポート430Aは、未使用中になると、ポート選択制御部440に対して、SM470上のサブバッファ領域に格納されたメッセージを含むデータ α のリード要求（換言すれば、図11に示したコマンドの発行）を、自発的に行うか、或いは、ポート選択制御部440からの定期的な問い合わせに応答して行う。

【0127】

ポート選択制御部440のパス切替制御部650は、そのリード要求の発行元のポート430Aとポート管理テーブル600から、そのポート430Aに対応した宛先ディスク側MP550のMP識別情報を把握し、そのMP識別情報をアドレスデコーダ451に通知する。

【0128】

アドレスデコーダ451は、ポート選択制御部440からMP識別情報を受けたときは、その識別情報を持つMP550に対応するスタックポインタを選択し、スタックポインタテーブル453に対し、そのMP550についてのリード要求を送信する。

【0129】

スタックポインタテーブル453は、そのMP550についてリード要求を受けたときは、リードポインタ及びライトポインタを1デクリメントし、更新後のポインタの値と、ポインタの更新対象となったMP550の識別情報とを、アドレス制御部456、バッファフル判定部454及び閾値判定部455に送信する。

【0130】

アドレス制御部456は、スタックポインタテーブル453から受けたリードポインタの値と、ポートコントローラ410内に予め設定されているSM470のベースアドレスとに基づいて、SM470のどこからデータをリードすべきかを示すSMリードアドレスを算出する。アドレス制御部456は、算出されたSMリードアドレスをメモリ制御部460に通知する（S27）。

【0131】

メモリ制御部460は、アドレス制御部456からSMリードアドレスを受けた場合、そのSMリードアドレスにあるデータ α （すなわち、サブバッファ領域に格納されているLMアドレス及びメッセージのセット）を読み出し（S28）、それを、ポート選択制御部440を介して、上記リード要求の発行元のターゲットポート430Aに渡す。それにより、そのデータ α が、ターゲットポート430Aの送信制御部432によって、所定フォーマットに変換されて（例えば、LMアドレス、コマンド種別として「LMライト」、及びライトデータを含んだコマンドに変換されて）、ディスクポート520Aの受信制御部516に送信される（S29、S30）。

【0132】

その後、上記ケース1と同様の流れで、ディスク側MP550宛のメッセージが、そのMP550Aを介在することなく、そのMP550AのLM560Aにライトされ、MP550Aの自分のLM560Aに対するポーリングにより、MP550Aは、自分宛のメッセージを取得することができる（S31～S34）。

【0133】

なお、この流れでは、スタックポインタテーブル453のライトポインタの値が更新される都度、更新後のライトポインタの値が閾値判定部455及びバッファフル判定部454に通知される。

【0134】

閾値判定部455は、スタックポインタテーブル453から通知されたライトポインタが第1閾値を超えたか否かを判定し、超えていれば、閾値オーバーが発生した旨と、それ

が発生した宛先MP 550Aの識別情報とをステータス生成部452に送信する。

【0135】

バッファフル判定部454は、スタックポインタテーブル453から通知されたライトポインタが第1閾値よりも大きい第2閾値を超えたか否かを判定し、超えていれば、バッファフルが発生した旨と、それが発生したMP 550Aの識別情報とをステータス生成部452に送信する。

【0136】

ステータス生成部452は、閾値判定部455から閾値オーバー発生の通知及びそれが発生したMP 550Aの識別情報を受けた場合、そのMP識別情報と閾値オーバーの発生を示すコードとを含んだステータス情報を生成する。また、ステータス生成部452は、バッファフル判定部454からバッファフル発生の通知及びそれが発生したMP 550Aの識別情報を受けた場合、そのMP識別情報とバッファフルの発生を示すコードとを含んだステータス情報を生成する。ステータス生成部452は、生成したステータス情報を、ポート選択制御部440及び所定のポート420（及び／又は430）を介して、所定のMP 360、550（例えば、全てのホスト側MP 360又はメッセージの送信元のホスト側MP 360A）に送信する。

【0137】

ステータス情報を受けたポート、例えばホストポート200Aの受信制御部319は、受信したステータス情報をステータス判定部317に渡す。

【0138】

ステータス判定部317は、受信制御部319から受けたステータス情報の宛先MP情報、バッファフル情報及びバッファ閾値オーバー情報に基づいて、SMバッファ閾値テーブル314及びSMバッファフルテーブル315を更新する。具体的には、例えば、ステータス判定部317は、バッファフル情報がバッファフルを示していれば、宛先MP情報が示すMP 550AについてのSMバッファフルテーブル315上のビットをオフからオフにする。同様に、例えば、ステータス判定部317は、バッファ閾値オーバー情報が閾値オーバーを示していれば、宛先MP情報が示すMP 550AについてのSMバッファ閾値テーブル314上のビットをオフからオフにする。

【0139】

SMバッファ閾値テーブル314は、或るMP 550Aに対応したビットがオンにされた場合、その旨を割り込み制御部313及びアクセス制御部316に通知する。

【0140】

また、SMバッファフルテーブル315は、或るMP 550Aに対応したビットがオンにされた場合、その旨を割り込み制御部313及びアクセス制御部316に通知する。

【0141】

割り込み制御部313は、或るディスク側MP 550Aについて閾値オーバーが発生した旨の通知をSMバッファ閾値テーブル314から受けた場合、この割り込み制御部313が搭載されているホストアダプタ300上にある全てのMP 360（又は、メッセージの送信元MP 360Aのみ）に正常割り込み信号を送信する。また、割り込み制御部313は、或るディスク側MP 550Aについてバッファフルが発生した旨の通知をSMバッファフルテーブル315から受けた場合、この割り込み制御部313が搭載されているホストアダプタ300上にある全てのMP 360（又は、メッセージの送信元MP 360Aのみ）に障害割り込み信号を送信する。

【0142】

正常割り込み信号を受けたMP 360Aのマイクロプログラムは、上記ディスク側MP 550についての閾値オーバーをリセットするための閾値オーバーリセット命令をリセット制御部312に発行する。その場合、リセット制御部312は、その閾値オーバーリセット命令に応答して、命令されたMP 550についてのSMバッファ閾値テーブル314のビットをオンからオフにする（すなわち、そのMP 550についての閾値オーバー状態を解除する）。SMバッファ閾値テーブル314は、ビットがオンからオフにリセットされ

た旨とそのリセットの対象となったMP識別情報を、アクセス制御部316に通知する。これにより、閾値オーバーが発生した宛先MP550Aに対して各MP360からメッセージが送信されても、アクセス制御部316を介して中継アダプタにそのメッセージが送信される。なお、MP360Aは、正常割込み信号を受けた場合は、閾値オーバーが発生したMP550Aを宛先とするメッセージの発行量又は発行頻度を減らすようにしても良い。

【0143】

障害割込み信号を受けたMP360のマイクロプログラムは、所定のイベントが発生するまで（例えば、バッファフルが解除された旨の通知をポートコントローラ410から受けるまで）、上記ディスク側MP550Aを宛先としたメッセージの発行頻度又は発行量を減らすか、或いは、直接LMライト方式から間接LMライト方式に切り替えて上記ディスク側MP550Aを宛先とするメッセージを発行し、所定のイベントが発生したら、通常通りに直接LMライト方式でその宛先にメッセージを発行する（マクロプログラムは、通常通りに直接LMライト方式のメッセージを発行し、アクセス制御部316が、そのメッセージを中継アダプタ400側へスルーすることを止めても良い）。

【0144】

なお、バッファフルがSMバッファフルテーブル315に登録された場合、そのバッファフルは例えば以下の方法で解除される。

【0145】

すなわち、バッファフル判定部454が、スタックポインタテーブル453から受けた更新後のライトポインタの値が、バッファフルの基準となる第2閾値よりも所定値小さい第3閾値に達した場合に、その旨をステータス生成部452に通知し、ステータス生成部452が、バッファフル解除を表すコードとバッファフル解除対象のMPの識別情報とを含んだステータス情報を送信する。そのステータス情報は、所定のルートを経由してホストハブ310のステータス判定部317が受信し、ステータス判定部317が、そのステータス情報の内容をリセット制御部312に通知する。リセット制御部312は、通知された内容から、バッファフル解除対象のMPの識別情報を特定し、そのMPについてのSMバッファフルテーブル315のビットをオンからオフにリセットすることで、そのMPについてのバッファフルを解除する。SMバッファフルテーブル315は、ビットがオンからオフにリセットされた旨とそのリセットの対象となったMP識別情報を、アクセス制御部316に通知する。これにより、例えば、バッファフルが発生したMP550を宛先としたメッセージがアクセス制御部316において止められていた場合に、それが解消されて、そのMP550を宛先とするメッセージがアクセス制御部316を介して中継アダプタに送信されるようになる。

【0146】

また、SMバッファフルテーブル315でバッファフルがリセットされた場合、その旨が割込み制御部313に通知され、割込み制御部313から、バッファフルのリセットに対応した種類の割込み信号が各ホスト側MP360に送信されても良い。その場合、その割込み信号を受けたホスト側MP360は、バッファフルが発生したMP550を宛先とするメッセージを間接LMライト方式（つまり共有メモリ方式）に従って送信していたならば、それを止めて、そのMP550に対して直接LMライト方式に従ってメッセージを送信することに戻しても良い。

【0147】

以下、これらについての具体例を説明する。

【0148】

図13は、アクセス制御部316がSMバッファフルテーブル315の情報に基づいてコマンドの送信を制御する場合の処理流れを示す。

【0149】

ホストアダプタ300上のアクセス制御部316は、SMバッファフルテーブル315の閾値登録代表ビットを参照し、それがオンかオフであるかを判定する（S51）。

【0150】

S51の結果、閾値登録代表ビットがオフであれば(S51でY)、どの宛先MPについてもバッファフルが生じておらず、直接ライトコマンドがSM470上のどのサブバッファ領域に蓄積されてもバッファフルが生じないことを意味するので、アクセス制御部316は、直接ライトコマンドが送信されるようにする。

【0151】

一方、S51の結果、閾値登録代表ビットがオンであれば(S51でN)、少なくとも1つの宛先MPについてバッファフルが生じていることを意味するので、アクセス制御部316は、SMバッファフルテーブル315において、コマンドの宛先MPに対応したビットがオンかオフかを判定する(S53)。S53の結果、アクセス制御部316は、そのビットがオフであれば、直接ライトコマンドが送信されるようにし(S52)、一方、そのビットがオンであれば、間接ライトコマンドが送信されるようにする(例えば、コマンド送信元MP360Aに対して、そのコマンドのLMアドレスをSMアドレスに代えた間接コマンドを送信させる)(S54)。

【0152】

図14は、SMバッファフルテーブル315でバッファフルがリセットされる場合の処理流れを示す。

【0153】

中継アダプタ400のバッファフル判定部454が、定期的に、バッファフルになったMPのライトポインタをリードする(S61)。

【0154】

S61の結果、そのライトポインタが、バッファフルの基準となる第2閾値よりも所定値小さい第3閾値以上でなければ(S61でN)、バッファフル判定部454は、バッファフルの解除の旨をステータス生成部452に通知する。それにより、上述した処理が行われて、ホストアダプタ300上のSMバッファフルテーブル315において、バッファフルのためオンであったビットがオフにされる(S62)。

【0155】

また、S61の結果、リードされたライトポインタが上記第3閾値以上であっても、バッファフルが発生してから一定時間長経過した場合には(S63でY)、バッファフル判定部454は、バッファフルの解除の旨をステータス生成部452に通知し、それにより、バッファフルのためオンであったビットがオフにされる(S62)。

【0156】

以上、上述した第1実施例によれば、送信側MPが発行したメッセージが、受信側MPを介在することなくそのMPが有するLMに直接ライトされる。受信側MPは、自分が持つLMをポーリングすることで、自分宛のメッセージを取得する。LMをリードする時間長は、SM470をリードする時間長に比べて短い(例えば10分の1以下程度である)。このため、受信側MPは、負担少なく、送信側MPから出力されたメッセージを取得することができる。

【0157】

具体的には、例えば、各MPの共有メモリへのポーリングの頻度を「 $5\mu\text{sec}$ 」に1回とし、共有メモリのパス数を「32」とすると、「32」個のMPがあると、MPや共有メモリにとって負担大となり、システムの性能の低下につながる。このとき、各MPと共有メモリ制御部間のパス(個別部)アクセス時間は800nsec、共有メモリ制御部と共有メモリ(共通部)のアクセス時間は200nsec、従って、個別パスが飽和するには、 $800\text{nsec} + 200\text{nsec} \times 32\text{MP} = 7200\text{nsec}$ となる。7200nsec以内の間隔で各MPが共有メモリをポーリングするとパスが飽和してしまい、アクセス時間は、急激に長くなりシステム性能を大幅に劣化させる。

【0158】

上述した第1実施例によれば、このような負担を軽減し、処理速度の高いディスクアレイ制御装置100を提供することが図れる。

【0159】

また、上述した第1実施例によれば、中継アダプタ400は、SM470を備え、そのSM470には、バッファ領域が用意される。直接LMライト方式のメッセージは、受信側アダプタに対応したポートの使用状況に応じて、そのバッファ領域に格納せずにそのアダプタに転送されたり、そのバッファ領域に一時的に格納されてから転送されたりする。すなわち、送信対象のメッセージが直接LMライト方式のメッセージの場合は、受信側アダプタに対応したポートの使用状況に応じて、SM470のバッファ領域を経由して受信側アダプタに転送される第1転送モードと、そのバッファ領域を経由せずにダイレクトに転送される第2転送モードとが選択的に行われる。これにより、送信側MP（例えばホスト側MP360）と受信側MP（例えばディスク側MP550）との間のメッセージの転送効率の向上が図れる。

【0160】

また、上述した第1実施例によれば、中継アダプタ400を介してメッセージを転送する際に、ホスト側ポート420とディスク側ポート430は独立して動作することが可能である。例えば、或るターゲットディスク側ポート430を経由すべきメッセージを或るホスト側ポート420が受信中でもあっても、そのターゲットディスク側ポート430は、ディスクハブ510に対して通信状態を維持する必要は無い。これにより、送信側MP（例えばホスト側MP360）と受信側MP（例えばディスク側MP550）との間のメッセージの転送効率の向上が図れる。

【0161】

なお、第1実施例では、下記のような幾つかの変形例が考えられる。

【0162】

例えば、各MP360は、アクセス制御部316に対し、SMアドレスとLMアドレスの双方を含んだコマンドを送信しても良い。その場合、アドレス制御部316は、各テーブル314、315の情報に基づき、直接LMライト方式を選択する場合は、受信したコマンドからSMアドレスを削除した新たなコマンドを生成して出力し、一方、間接LMライト方式を選択する場合は、上記受信したコマンドからLMアドレスを削除した（或いは削除せずにそのまま残した）コマンドを出力するようにしても良い。

【0163】

また、例えば、ホスト側ポート420A～420n及びディスク側ポート430A～430nの各々は、LMアドレス範囲に代えて又は加えて、宛先MP識別情報に対応付けられ、その対応関係がポート管理テーブル600に登録されても良い。また、直接ライトコマンド及び間接ライトコマンドには、各MPが記憶するアドレスマップ4から把握可能な宛先MP識別情報が含まれても良い。すなわち、中継アダプタ400では、受信したコマンドを転送する際、そのコマンドに含まれている宛先MP識別情報から、出力先のポートが選択されても良い。

【0164】

また、例えば、各アダプタ300、400及び／又は500上にある1以上のポートは、物理的に分離していても良いし、物理的には分離していなくても論理的に分離していても良い。同様に、同一ポート内にある送信制御部及び受信制御部も、物理的に分離していても良いし、物理的には分離していなくても論理的に分離していても良い。

【実施例2】

【0165】

上述した第1実施例では、コマンドを転送する中継アダプタがマスタ、そのコマンドを受ける受信側アダプタ（例えばディスクアダプタ500）がスレーブとなる中継部マスタ方式でコマンド転送が行われるが、これから説明する第2実施例では、中継アダプタがスレーブ、受信側アダプタがマスタとなる受信側マスタ方式でコマンド転送が行われる。以下、詳述する。

【0166】

図15は、本発明の第2実施例におけるディスクアダプタの構成例である。なお、この

図において、図5に示したディスクアダプタ500と同様の機能を有するものについては同一の参照番号を付す。また、以下の説明では、第1実施例と重複した部分については、説明を簡略或いは省略する。

【0167】

この第2実施例に係るディスクアダプタ900の各ディスクポート901A、901Bには、中継アダプタからパス割込み信号を受けるパス割込制御部903が備えられる。

【0168】

パス割込制御部903は、パス割込み信号の立ち下がり（或いは立ち上がり）を検出した場合に、同一ポート901A又は901B内の送信制御部904に、パス割込み有り（換言すれば、直接ライトコマンド有り）を通知する。

【0169】

送信制御部904は、パス割込み有りを受けた場合、中継アダプタ400に、図16に例示するようなリード割込みコマンド、すなわち、割込みアドレス（「INTアドレス」と図示）とコマンド種別として「共有メモリリード」とを含んだコマンドを送信する。それにより、受信制御部516が、中継アダプタ400から、図17に示すようなデータ、すなわち、LMアドレスとライトデータとを含んだデータを受信する。なお、INTアドレスは、SM470のサブバッファ領域に格納されたデータ又はポート選択制御部440のバッファ内に格納されたデータの格納場所を割出すことができるものであれば、どのようなものでも良い（例えば、各ポートに固有に割り当てられたアドレスであっても良い）。

【0170】

図18は、第2実施例において、SM470上の所定サブバッファ領域に直接ライトコマンドが格納される場合の、直接ライトコマンドのやり取りの流れを示す。図19は、その場合のコマンド送受信のタイミングチャートを示す。

【0171】

中継アダプタの或るホスト側ポート420Aが、直接ライトコマンド（LMアドレス、コマンド種別「LMライト」、ライトデータを含んだコマンド）を受信したら（S71、S72）、そのコマンドに含まれているLMアドレスとポート管理テーブル600とに基づいてターゲットポート430Aを識別し、そのターゲットポート430Aに対して所定のリクエスト信号を送信する（S73）。また、ポートコントローラ410内では、受信した直接ライトコマンド内のLMアドレスに基づいてSMアドレスが算出され、そのSMアドレスと、データ α （LMアドレス及びライトデータ）とがメモリ制御部460に取込まれ（S74）、メモリ制御部460によって、データ α がSM470上の所定サブバッファ領域に格納される。

【0172】

また、ホスト側ポート420Aは、直接ライトコマンドを受信した旨のステータスを接続先ホストポート200Aに送信する（S75）。

【0173】

ターゲットポート430Aは、ホスト側ポート420Aからのリクエスト信号に応答して、接続先ディスクポート901Aのパス割込制御部903へのパス割込み信号を立ち上げる（S76）。そして、ターゲットポート430Aは、上記データ α がSM470に格納された後、そのパス割込み信号を立ち下げる。

【0174】

パス割込制御部903は、パス割込み信号の立下りを検出したら、パス割込み有りを送信制御部904に通知する。

【0175】

パス割込み有りが通知された送信制御部904は、上述したリード割込みコマンド（INTアドレスとコマンド種別「共有メモリリード」とを含んだコマンド）を、ターゲットポート430Aに送信する（S77）。その結果、リード割込みコマンド内のINTアドレスから一意にSMアドレスが識別され（例えば1以上のINTアドレスと1以上のSM

アドレスとが対応付けられたテーブルから所定SMアドレスが識別され)、識別されたSMアドレスからデータ α が読み出され(S78、S79)、読み出されたデータ α が、ディスクアダプタ900に送信される(S80)。そして、そのデータ α 内のLMアドレスに、そのデータ α 内のライトデータが直接ライトされ(S81)、宛先MP550Aが自分のLM560の所定LMアドレスをポーリングすることにより、そのライトされたライトデータを取得することができる(S82～S85)。

【0176】

以上が、第2実施例についての説明である。なお、この第2実施例では、直接ライトコマンドは、必ずSM470のサブバッファ領域に一旦格納されてから宛先MP側に転送されても良いし、格納されずに宛先MP側に転送されても良い。後者の場合は、例えば以下の流れで転送が行われても良い。すなわち、中継アダプタ400において直接ライトコマンドの受信後、ターゲットポート430Aが、接続先ディスクポート901Aのパス割込制御部903へのパス割込み信号を立ち上げ、その時点から所定時間経過後、そのパス割込み信号を立ち下げる。その場合、送信制御部904が、上述したリード割込みコマンドを、ターゲットポート430Aに送信し、それに応答して、受信制御部516が、ポート選択制御部440内の所定バッファから読み出された直接ライトコマンドを受信しても良い。

【0177】

また、例えば、図15では、ディスクアダプタを示したが、メッセージは、ディスクアダプタからホストアダプタに転送される場合もあるので、ホストアダプタの各ポートも図15に示したディスクアダプタの各ポートと同様の構成になっていても良い。

【実施例3】

【0178】

次に、本発明の第3実施例について説明する。

【0179】

この第3実施例では、各MP(例えばホストMP360A)は、直接LMライト方式や間接LMライト方式でリードコマンドを所望格納先LM(例えばディスク側LM550A)に格納することで、そのLMからリードされたメッセージを受ける。

【0180】

この第3実施例は、上述した第1実施例の応用例である。すなわち、リードコマンドのコマンド送信側MPが、直接LMライト方式又は間接LMライト方式でコマンド受信側MPのコマンド受信側LMにそのリードコマンドをライトすると、コマンド受信側MPが、そのリードコマンドに応答して、コマンド受信側LMから、コマンド送信側MPを宛先とするメッセージを取得し、そのメッセージを、直接LMライト方式又は間接LMライト方式で、コマンド送信側MPのコマンド送信側LMに格納する。その後、コマンド送信側MPが、そのコマンド送信側LMをポーリングしてメッセージを取得する。これにより、コマンド送信側MPが、コマンド受信側LMからメッセージを取得したことになる。

【0181】

すなわち、第1実施例では、メッセージ送信元MPは、リードコマンドを受けることなく能動的にメッセージを所定格納先LMに発行するが、第3実施例では、自LMにメッセージを格納しておき、他MPからリードコマンドを受けたときに、それに応答して、自LMからメッセージを読み出し、それを、そのコマンド送信元MPのLMに送信する。

【0182】

以下、この第3実施例について詳細に説明する。なお、以下の説明において、第1実施例と重複する部分については説明を省略又は簡略する。

【0183】

図20は、第3実施例において、各MP360(及び550)に格納されているアドレスマップの一例を示す。

【0184】

各MP360(及び550)、例えば、MP識別情報(「MP#」と図示)が「00」

であるホストMP 3 6 0 に格納されているアドレスマップ1 2 には、種々の情報が記録されているが、主要な情報の例として、第1 実施例で説明した送信時SMアドレス情報及び受信時SMアドレス情報の他に、自MPのLMリード／ライトアドレス情報、他MPのLMライトアドレス情報及び他MPのLMリードアドレス情報がある。

【0 1 8 5】

自MPのLMリード／ライトアドレス情報は、自分が有する1 以上のLM 3 5 0（又は5 6 0）におけるLMアドレス範囲であって、自MP 3 6 0 に対するリードコマンドが格納されるLMリードアドレス範囲（1 又は複数のLMアドレス）と、他MP 5 5 0 に対するメッセージの格納先となるLMライトアドレス範囲（1 又は複数のLMアドレス）とを含んだ情報である。LMリードアドレス範囲もLMライトアドレス範囲も、各他MP 毎（及び／又は各他LM 毎）に用意される（例えば、各他MP のMP # に対応付けられている）。すなわち、1 以上の他MP にそれぞれ対応した1 以上のLMリードアドレス範囲の各々は、それに対応した他MP にとっては、その他MP が発行するリードコマンドの格納先LMライトアドレス範囲である。また、1 以上の他MP にそれぞれ対応した1 以上のLMライトアドレス範囲の各々は、それに対応した他MP にとっては、その他MP のメッセージのリード先となるLMリードアドレス範囲である。

【0 1 8 6】

他MPのLMライトアドレス情報は、他MP が有する他LM におけるLMアドレス範囲であって、自MP 3 6 0 が発行したリードコマンドの格納先LMライトアドレス範囲（1 又は複数のLMアドレス）を表す情報である。このLMライトアドレス範囲は、図示のように、各他MP 毎に用意される（例えば、各他MP のMP # に対応付けられている）。すなわち、1 以上の他MP にそれぞれ対応した1 以上のLMライトアドレス範囲の各々は、それに対応した他MP にとっては、その他MP のポーリングによるリード先となるLMリードアドレス範囲である。

【0 1 8 7】

他MPのLMリードアドレス情報は、他MP が有する他LM におけるLMアドレス範囲であって、自MP 3 6 0 のリード先であるLMリードアドレス範囲（1 又は複数のLMアドレス）を表す情報である。このLMリードアドレス範囲は、図示のように、各他MP 毎に用意される（例えば、各他MP のMP # に対応付けられている）。すなわち、1 以上の他MP にそれぞれ対応した1 以上のLMリードアドレス範囲の各々は、それに対応した他MP にとっては、その他MP の上記自MP に対するメッセージの格納先となるLMライトアドレス範囲である。

【0 1 8 8】

以上のように、各MP 3 6 0、5 5 0 のプログラムアドレス空間に、上記のようなアドレスマップ1 2 を記録しておく。これにより、そのMP 3 6 0（又は5 5 0）のマイクロプログラムの処理によって、任意のターゲットMP 5 5 0（又は3 6 0）のLMライトアドレスに、任意のLMリードアドレスを指定したリードコマンドが任意のLMライトアドレスにライトされる。

【0 1 8 9】

以下、この第3 実施例において、コマンド送信側MP がホスト側MP 3 6 0 A、コマンド受信側MP がディスク側MP 5 5 0 Aである場合を例に採り、直接LMリード方式でメッセージが読み出される場合の流れについて、以下の2 つのケースに場合分けして説明する。

【0 1 9 0】

(1) 第1 ケース：直接LMライト方式のリードコマンドがSM 4 7 0 のサブバッファ領域に一時格納されることが無い場合。

【0 1 9 1】

この第1 ケースにおいてアダプタ3 0 0、4 0 0 及び5 0 0 が行うデータのやり取りの流れを図2 1 に示す。

【0 1 9 2】

ホスト側MP 360Aは、自分が記憶するアドレスマップ12から、所望相手MP 550Aに対応したLMリードアドレス（どこからデータを読むかを示すアドレス）及びLMライトアドレス（リードコマンドの格納先）を取得して、取得したLMリードアドレスを指定するリードコマンドと、取得したLMライトアドレスとを含んだリード命令をMPバス2に出力する。すると、そのリード命令は、MPバス制御部311及びアクセス制御部316を介して、自MP 360Aに対応したホストポート200Aの送信制御部318に送信される。送信制御部318は、そのリード命令を所定のフォーマットに変換し、フォーマット変換後のコマンド群（すなわち、LMライトアドレスと、コマンド種別「直接LMリード」を表す上記リードコマンド）を、ホストポート200Aの接続先のホスト側ポート420Aが有する受信制御部421に送信する（S102）。

【0193】

ポートコントローラ410のホスト側ポート420Aが有する受信制御部421は、ホストポート200Aの送信制御部318から受信したコマンド群を、ポート選択制御部440内の所定バッファ601Aに一時格納する。また、その受信制御部421は、そのコマンド群に含まれているLMライトアドレスとポート管理テーブル600とに基づいて、どのディスク側ポート430A～430nを使用すべきかを識別し、識別されたターゲットポート430A（つまり、格納先LMアドレスに対応したディスク側ポート）に対して、所定のリクエスト信号を送信する（S103）。ホスト側ポート420Aが、そのリクエスト信号に応答して、ターゲットポート430A（又はそれに対応したポートバス制御部602A）から使用中で無い旨を表す信号（「not busy」と図示）を受けた場合（S104）、すなわち、ターゲットポート430が未使用である場合、バッファ601A内のコマンド群が読み出されてターゲットポート430Aの送信制御部432を介してディスクアダプタ500に転送される（S105、S106）。また、ホスト側ポート420Aの送信制御部442は、コマンド群の転送が完了したことを表す通知を、接続先ホストポート200Aの受信制御部319に送信する（それにより、ホストアダプタ300において、出力したコマンド群の転送が終了したことを認識することができる）（S107）。

【0194】

中継アダプタ400からコマンド群を受けたディスクポート520Aの受信制御部516は、そのコマンド群（すなわち、LMライトアドレス及びリードコマンド）をLMアドレス制御部514に送る。

【0195】

LMアドレス制御部514は、受信制御部516から受けたLMライトアドレスと、LMベースアドレス設定部513が記憶しているベースアドレスとを用いて正確なLMアドレスを算出し、そのLMライトアドレスと、受信制御部516から受けたリードコマンドとをMPバス制御部515に渡す。

【0196】

MPバス制御部515は、プロセッサバス6を介して、LMアドレス制御部514から受けたLMライトアドレス（換言すれば、ディスク側MP 550Aが記憶しているアドレスマップの自MPのLMリード／ライトアドレス情報における、LM 560Aのうち送信元MP 360に対応したLMリードアドレス）に、そのLMアドレス制御部514から受けたリードコマンドを書き込む（S108）。

【0197】

ディスク側MP 550Aは、例えば、自LM 560Aにおける自MPライトアドレス情報のうち、送信元MP 360に対応したLMライトアドレスに、送信元MP 360に対するメッセージを書き込んでおく。

【0198】

ディスク側MP 550Aは、自分が記憶しているアドレスマップの自MPのLMリード／ライトアドレス情報に基づいて、自分が有する自LM 560Aのうち送信元MP 360に対応したLMリードアドレスをポーリングする（すなわち、そのLMリードアドレスを

定期的にリードする)。その場合、そのLMリードアドレスにリードコマンドがあれば、ディスク側MP550は、そのリードコマンドに指定されるLMリードアドレスからメッセージを取得し、そのメッセージを、直接LMライト方式又は間接LMライト方式で、リードコマンド送信元であるホスト側MP360Aに送信する(ここで取得され送信されるメッセージは、アドレスマップの構成により、必ず、リードコマンドの送信元MPに対するメッセージになっている)。なお、そのメッセージがどのような流れでホスト側LM350Aに格納されるのかということについては説明を省略するが、それについては、第1実施例の説明から理解することができる。

【0199】

(2) 第2ケース：直接LMライト方式のリードコマンドが送信される場合であって、そのリードコマンドがSM470のサブバッファ領域に一時格納される場合。

【0200】

この第2ケースにおいてアダプタ300、400及び500が行うデータのやり取りの流れを図22に示す。

【0201】

ホスト側ポート420Aの受信制御部421は、上記第1ケースと同様の流れで上記コマンド群(LMライトアドレス及びリードコマンド)を受信して(S121、S122)、所定バッファ601Aに一時格納し、そのコマンドに含まれているLMアドレスとポート管理テーブル600とに基づいてターゲットポート430Aを識別し、そのターゲットポート430Aに対して所定のリクエスト信号を送信する(S123)。ホスト側ポート420Aは、そのリクエスト信号に応答して、ターゲットポート430Aから使用中である旨を表す信号(「busy」と図示)を受けた場合(S124)、すなわち、ターゲットポート430が使用中である場合、バス切替制御部650が、受信したコマンド中のLMライトアドレスをアドレスデコーダ451に通知し、且つ、受信したコマンド群に含まれているLMライトアドレス及びリードコマンドをデータ α としてメモリ制御部460に送信する(S125)。その後、ホスト側ポート420Aは、コマンド群の転送が完了したことを表す通知を、ホストポート200Aの受信制御部319に送信する(S126)。

【0202】

アドレスデコーダ451は、ポート選択制御部440からLMライトアドレスが通知された場合、そのLMライトアドレスからリードコマンドの宛先MP550Aを識別して、識別されたMP550Aに対応するスタックポインタを選択し、スタックポインタテーブル453に対し、そのMP550Aについてのライト要求を送信する。

【0203】

スタックポインタテーブル453は、アドレスデコーダ451から或る宛先MP550Aについてライト要求を受けたときは、そのMP550Aのライトポインタを1インクリメントし、更新後のライトポインタの値と、ポインタの更新対象となったMP550の識別情報とを、アドレス制御部456、バッファフル判定部454及び閾値判定部455に送信する。

【0204】

アドレス制御部456は、スタックポインタテーブル453から受けるライトポインタの値と、ポートコントローラ410内に予め設定されているSM470のベースアドレスとに基づいて、SM470のどこにデータを格納すべきかを示すSMライトアドレスを算出する。アドレス制御部456は、算出されたSMライトアドレスをメモリ制御部460に通知する。

【0205】

メモリ制御部460は、アドレス制御部456からSMライトアドレスを受けた場合、そのSMライトアドレスに、ポート選択制御部440から受けたデータ α 、すなわち、LMライトアドレス及びリードコマンドを格納する。

【0206】

その後、ターゲットポート430Aは、未使用中になると、ポート選択制御部440に

対して、SM470上のサブバッファ領域に格納されたメッセージを含むデータ α のリード要求（換言すれば、図11に示したコマンドの発行）を、自発的に行うか、或いは、ポート選択制御部440からの定期的な問い合わせに応答して行う。

【0207】

ポート選択制御部440のパス切替制御部650は、そのリード要求の発行元のポート430Aとポート管理テーブル600から、そのポート430Aに対応した宛先ディスク側MP550AのMP識別情報を把握し、そのMP識別情報をアドレスデコーダ451に通知する。

【0208】

アドレスデコーダ451は、ポート選択制御部440からMP識別情報を受けたときは、その識別情報を持つMP550Aに対応するスタックポインタを選択し、スタックポインタテーブル453に対し、そのMP550Aについてのリード要求を送信する。

【0209】

スタックポインタテーブル453は、そのMP550Aについてリード要求を受けたときは、リードポインタ及びライトポインタを1デクリメントし、更新後のポインタの値と、ポインタの更新対象となったMP550Aの識別情報とを、アドレス制御部456、バッファフル判定部454及び閾値判定部455に送信する。

【0210】

アドレス制御部456は、スタックポインタテーブル453から受けたリードポインタの値と、ポートコントローラ410内に予め設定されているSM470のベースアドレスとに基づいて、SM470のどこからデータをリードすべきかを示すSMリードアドレスを算出する。アドレス制御部456は、算出されたSMリードアドレスをメモリ制御部460に通知する（S127）。

【0211】

メモリ制御部460は、アドレス制御部456からSMリードアドレスを受けた場合、そのSMリードアドレスにあるデータ α （すなわち、サブバッファ領域に格納されているLMライトアドレス及びメッセージのセット）を読み出し（S128）、それを、ポート選択制御部440を介して、上記リード要求の発行元のターゲットポート430Aに渡す。それにより、そのデータ α が、ターゲットポート430Aの送信制御部432によって、所定フォーマットに変換されて（例えば、LMライトアドレス及びリードコマンドを含んだコマンド群に変換されて）、ディスクポート520Aの受信制御部516に送信される（S129）。

【0212】

その後、上記ケース1と同様の流れで、ディスク側MP550A宛のリードコマンドが、そのMP550Aを介在することなく、そのMP550AのLM560Aにライトされ（S130）、MP550Aの自分のLM560Aに対するポーリングにより、MP550Aは、自分宛のリードコマンドを取得することができる。なお、その後は、上記第1ケースと同様の流れで、そのリードコマンドから認識されるLMリードアドレス内のメッセージが、リードコマンドの送信元MP360AのLM350Aに格納される。

【0213】

以上が、本発明の第3実施例についての説明である。

【0214】

なお、この第3実施例において、もし、コマンド受信側LMにおいて、コマンド送信側MPから発行されたリードコマンドから認識されるLMリードアドレスにメッセージが存在しない場合には、そのLMリードアドレスにメッセージが格納された後、直ちにメッセージが読み出されてコマンド送信側LMに向けて送信されても良い。

【0215】

また、説明は省略したが、SMバッファ閾値テーブル314やSMバッファフルテーブル315に対する監視は、リードコマンドがSM470に格納された場合にも同様に行なわれても良い。すなわち、この第3の実施例でも、その監視結果に基づいて所定のステータス

タスがコマンド送信側アダプタに発行され、そのステータスに基づいて、コマンド送信側アダプタにおいて、リードコマンドの発行頻度や発行量が制限される等の処理を行うことが可能である。

【0216】

また、第3実施例において、例えば、アドレスマップ12は、どの自LMリードアドレスにリードコマンドが格納されたときにはどの自LMライトアドレスに格納したメッセージを返信するかを、そのアドレスマップ12を記憶するMPが認識することができるように作られていても良い。その場合は、リードコマンドに、LMリードアドレスは含まれても良い。

【0217】

また、第3実施例において、コマンド送信側MPからコマンド受信側LMにリードコマンドが格納されること、及び／又は、コマンド受信側LMから読み出された、コマンド送信側MPにとってのリード情報（コマンド送信先にとってはライト情報）がコマンド送信側LMに格納されることは、上述した中継部マスタ方式で行われても良いし、第2実施例で説明した受信側マスタ方式で行われても良い。

【0218】

以上、本発明の好適な幾つかの実施例を説明したが、これらは本発明の説明のための例示であって、本発明の範囲をこれらの実施例にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。

【0219】

例えば、第1～3実施例では、ホストコンピュータ200が出力したライト対象データ（例えば、ユーザに指定されたデータファイル）が、ホストアダプタ300、中継アダプタ400及びディスクアダプタ500を介して、物理ディスク群9上の物理的又は論理的な記憶領域に書き込まれたり、その記憶領域からリードされたリード対象データが、ディスクアダプタ500、中継アダプタ400及びホストアダプタ300を介して（厳密には、SM470、又は、ディスクアダプタ500とホストアダプタ300との間に備えられる別のメモリ（例えばキャッシュメモリ）を介して）、ホストコンピュータ200に送信されたりするが、LMに直接ライトされる情報は、MP間で受け渡しされる上記メッセージに限らず、上記ライト対象データやリード対象データであっても良い。その場合、通常、そのライト対象データやリード対象データのデータサイズはデータ毎に異なるが、MP間で受け渡しされる際には、所定サイズ毎に分割されてから受け渡しされても良い（所定サイズに満たない場合には、ダミーデータを追加して所定サイズにされても良い）。また、その際、例えば、直接LMライト方式として、上記別のメモリ（例えばキャッシュメモリ）を介さずにライト対象データ又はリード対象データが受け渡しされても良いし、関節LMライト方式として、上記別のメモリ（例えばキャッシュメモリ）に一時ライト対象データ又はリード対象データを格納した上で、ライト対象データ又はリード対象データが受け渡しされても良い。

【0220】

また、例えば、上述したホストアダプタ300とディスクアダプタ500とをどのようにに接続するかには種々の接続バリエーションが考えられる。例えば、第1の接続バリエーションでは、図23に示すように、分離したモジュール（例えばディスクアレイ制御装置）1000、910、920にあるアダプタ300と500とが、中継アダプタ400を介して接続される。また、例えば、第2の接続バリエーションでは、図24に示すように、n枚のアダプタ300又は500を有する分離した複数のモジュールが、中継装置960を介して通信可能に接続される。また、例えば、第3の接続バリエーションでは、図25に示すように、1以上のホストアダプタ300と中継アダプタ400とを有する第1モジュール（例えばディスクアレイ制御装置）と、1以上のディスクアダプタ500と中継アダプタ400とを有する第2モジュール（例えばディスクアレイ制御装置）とがあり、複数のモジュールにそれぞれ搭載の複数の中継アダプタ400同士が接続される。

【0221】

また、例えば、ディスクアレイ装置 100 では、第 1 実施例のように、各 MP が能動的に他 MP にライト情報（例えば、メッセージ、リード対象データ又はライト対象データ）を送信する第 1 モードと、第 3 実施例のように、各 MP が他 MP からのリードコマンドに应答してライト情報を送信する第 2 モードとが選択的に行われるようにしても良い。それら 2 種類のモードのうちのいずれを採用するかは、手動で又は自動で、所定タイミングに設定されても良い。

【図面の簡単な説明】

【0222】

【図 1】 本発明の第一実施例に係るディスクアレイ制御装置の構成を示す。

【図 2】 各 MP 360（及び 550）に格納されているアドレスマップの一例を示す。

【図 3】 ホストアダプタ 300 の機能及び構成を詳細に示すブロック図である。

【図 4】 中継アダプタ 400 の機能及び構成を詳細に示すブロック図である。

【図 5】 ディスクアダプタ 500 の機能及び構成を詳細に示すブロック図である。

【図 6】 ホストハブ 310 からポートコントローラ 410 に送られるコマンドの構成例を示す。

【図 7】 ポートコントローラ 410 からホストハブ 310 が受信するステータス情報の構成例を示す。

【図 8】 ポートコントローラ 410 からディスクハブ 510 が受信するコマンドの構成例を示す。

【図 9】 ポート選択制御部 440 の概要を示す。

【図 10】 ポート選択制御部 440 の具体例を示す。

【図 11】 直接 LM ライト方式のメッセージが送信される場合であって、そのメッセージが SM 470 のサブバッファ領域に一時格納されることが無い場合に、アダプタ 300、400 及び 500 が行うデータのやり取りの流れを示す。

【図 12】 直接 LM ライト方式のメッセージが送信される場合であって、そのメッセージが SM 470 のサブバッファ領域に一時格納される場合に、アダプタ 300、400 及び 500 が行うデータのやり取りの流れを示す。

【図 13】 アクセス制御部 316 が SM バッファフルテーブル 315 の情報に基づいてコマンドの送信を制御する場合の処理流れを示す。

【図 14】 SM バッファフルテーブル 315 でバッファフルがリセットされる場合の処理流れを示す。

【図 15】 本発明の第 2 実施例におけるディスクアダプタの構成例である。

【図 16】 リード割込みコマンドの構成例を示す。

【図 17】 受信側アダプタからのリード割込みコマンドに应答して中継アダプタから出力されるデータの構成例を示す。

【図 18】 第 2 実施例において、SM 470 上の所定サブバッファ領域に直接ライトコマンドが格納される場合の、直接ライトコマンドのやり取りの流れを示す。

【図 19】 図 18 におけるコマンド送受信のタイミングチャートを示す。

【図 20】 本発明の第 3 実施例におけるアドレスマップの一例。

【図 21】 第 3 実施例において、直接 LM ライト方式のリードコマンドが SM 470 のサブバッファ領域に一時格納されることが無い場合に行なわれるデータのやり取りの流れを示す。

【図 22】 第 3 実施例において、直接 LM ライト方式のリードコマンドが送信される場合であって、そのリードコマンドが SM 470 のサブバッファ領域に一時格納される場合に行なわれるデータのやり取りの流れを示す図。

【図 23】 ホストアダプタ 300 とディスクアダプタ 500 の第 1 の接続バリエーションの一例を示す図。

【図 24】 ホストアダプタ 300 とディスクアダプタ 500 の第 2 の接続バリエーションの一例を示す図。

【図 2 5】 ホストアダプタ 3 0 0 とディスクアダプタ 5 0 0 の第 3 の接続バリエーションの一例を示す図。

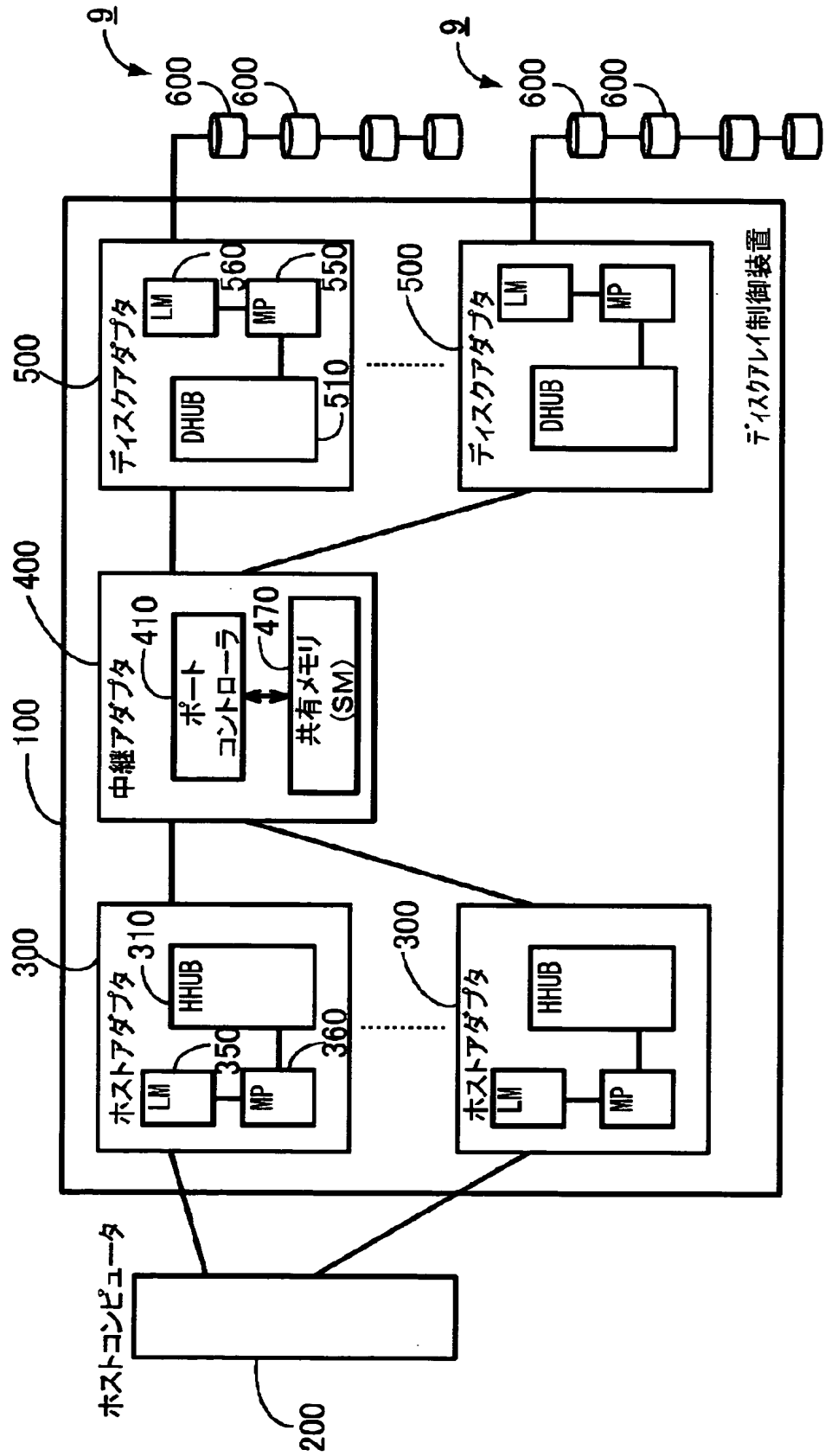
【符号の説明】

【 0 2 2 3 】

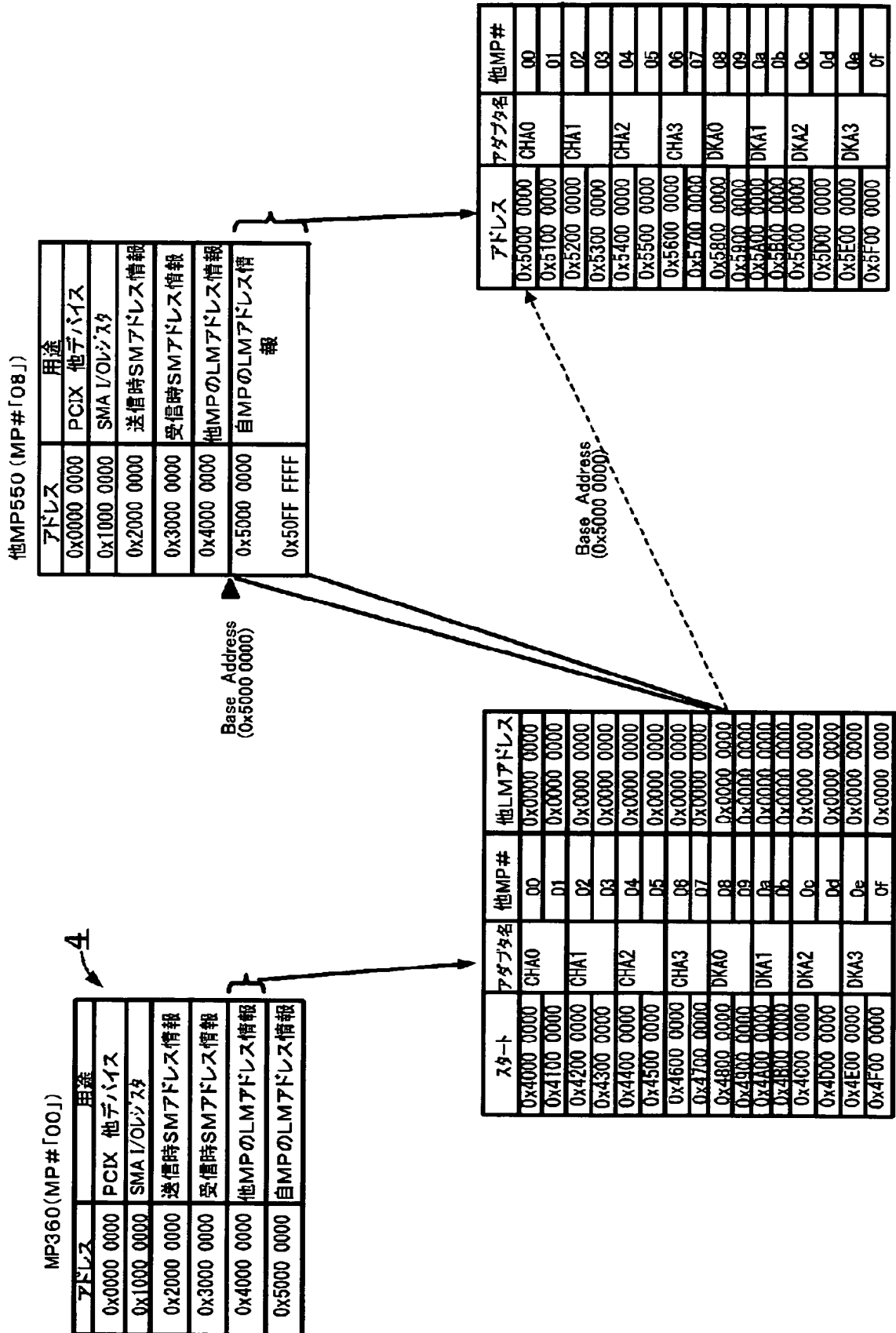
- 1 0 0 ディスクアレイ制御装置
- 2 0 0 ホストコンピュータ
- 6 0 0 磁気ディスク装置
- 3 0 0 ホストアダプタ
- 3 6 0 ホスト側マイクロプロセッサ
- 3 5 0 ホスト側ローカルメモリ
- 3 1 0 ホストハブ
- 5 0 0 ディスクアダプタ
- 5 5 0 ディスク側マイクロプロセッサ
- 5 6 0 ディスク側ローカルメモリ
- 5 1 0 ディスクハブ
- 4 0 0 中継アダプタ
- 4 1 0 ポートコントローラ
- 4 7 0 共有メモリ

●
【書類名】 図面

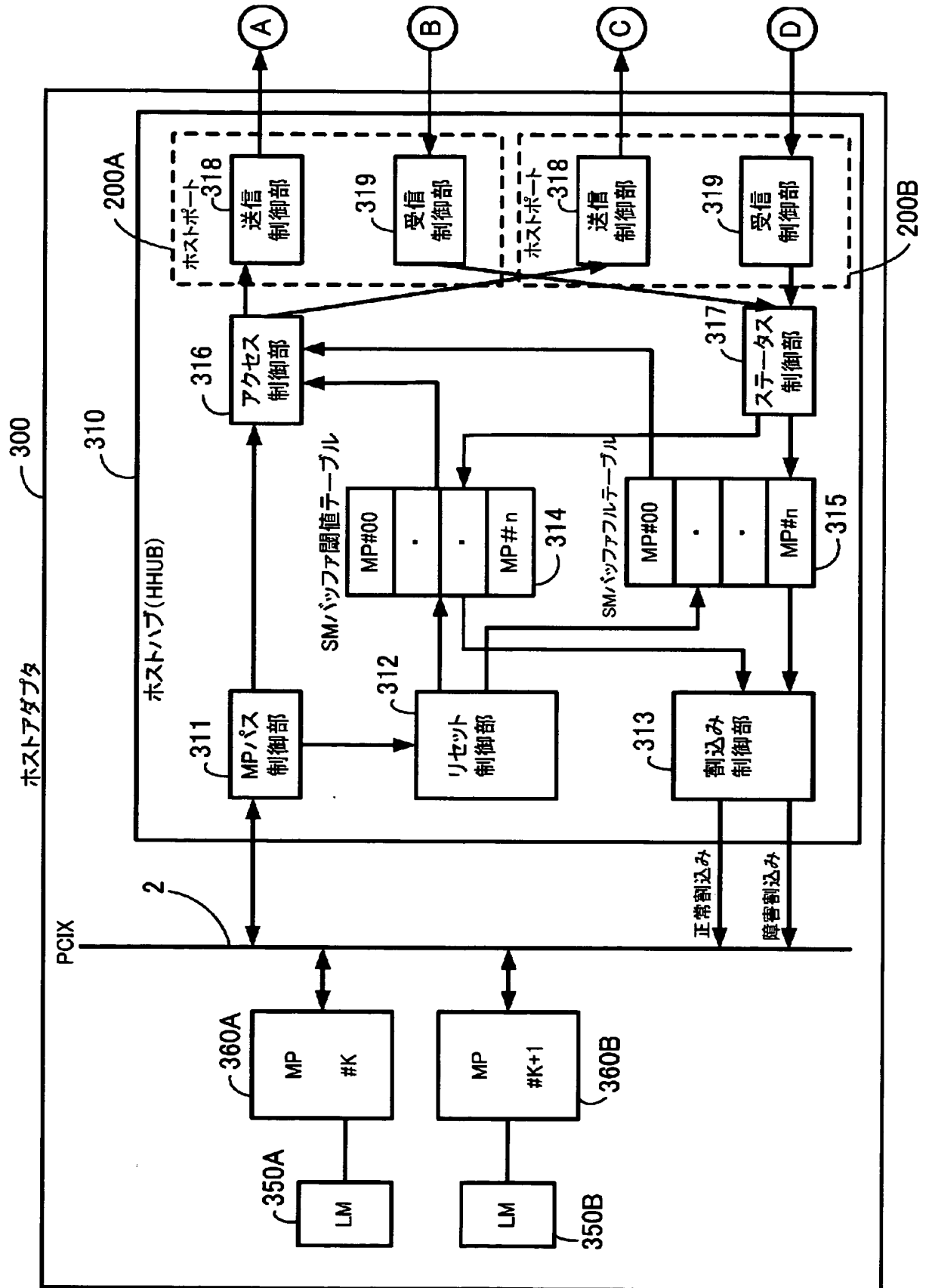
【図 1】



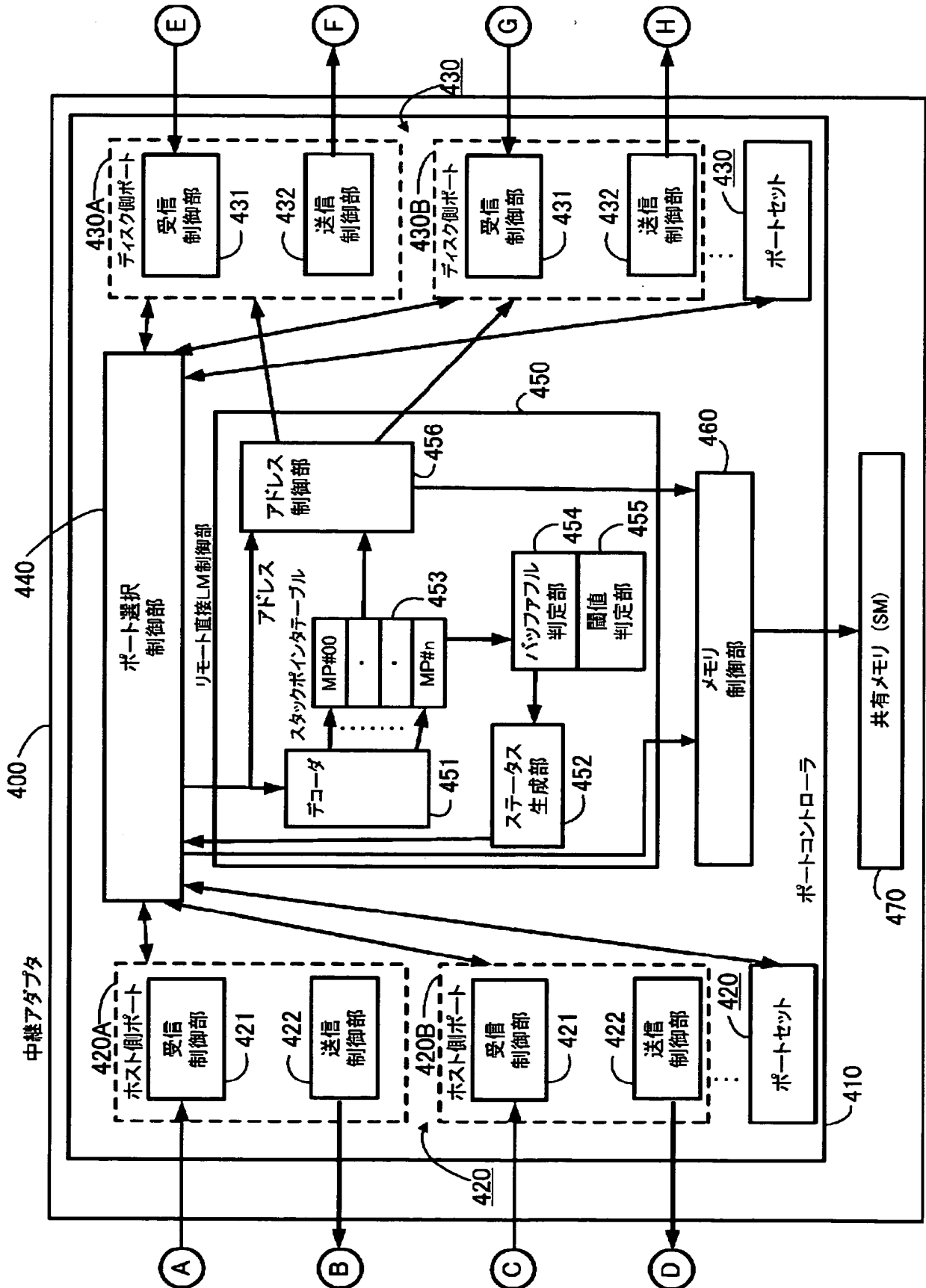
【図 2】



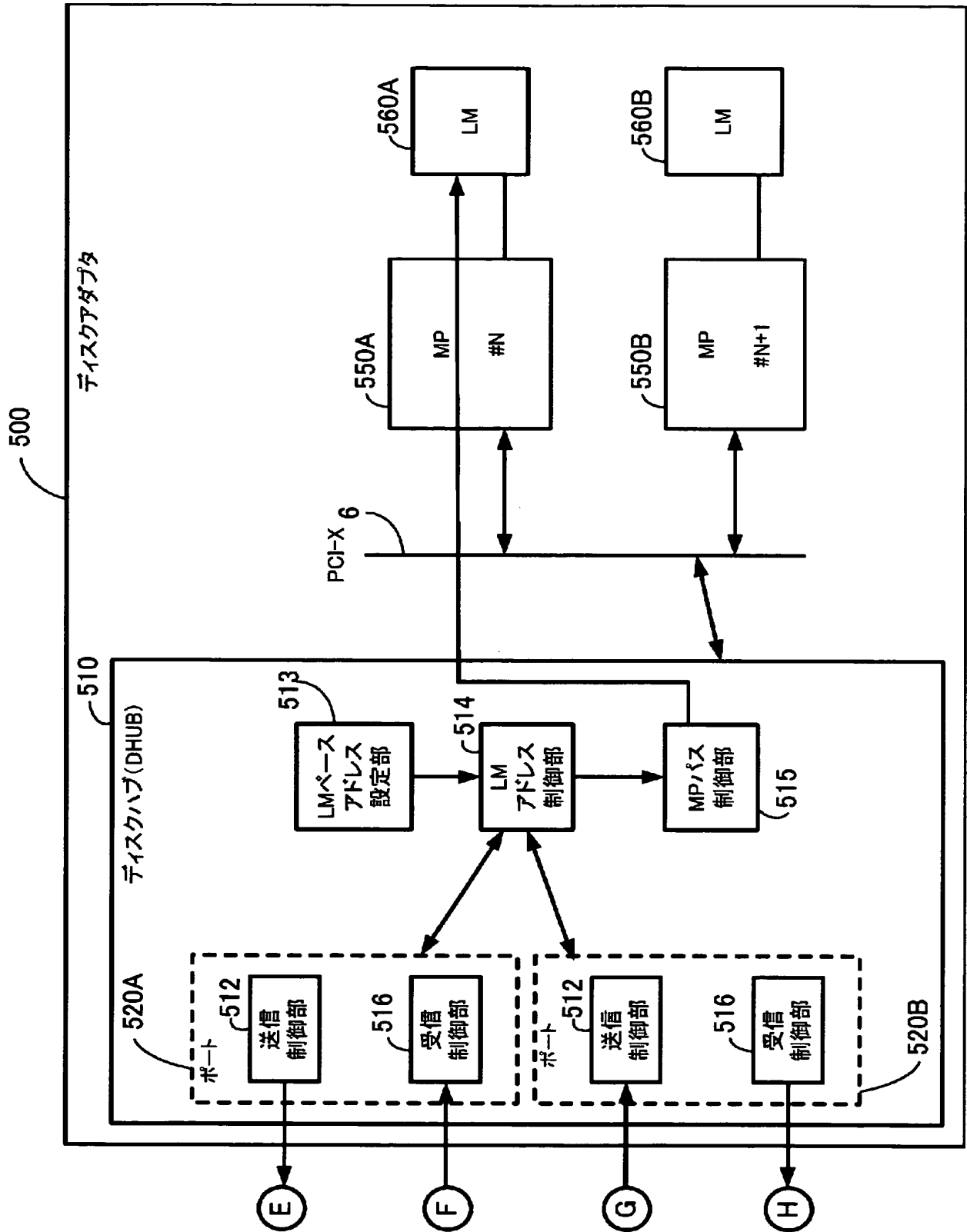
【図 3】



【図 4】



【図 5】



【図 6】

アドレス情報	コマンド種別	ライトデータ	CRC
--------	--------	--------	-----

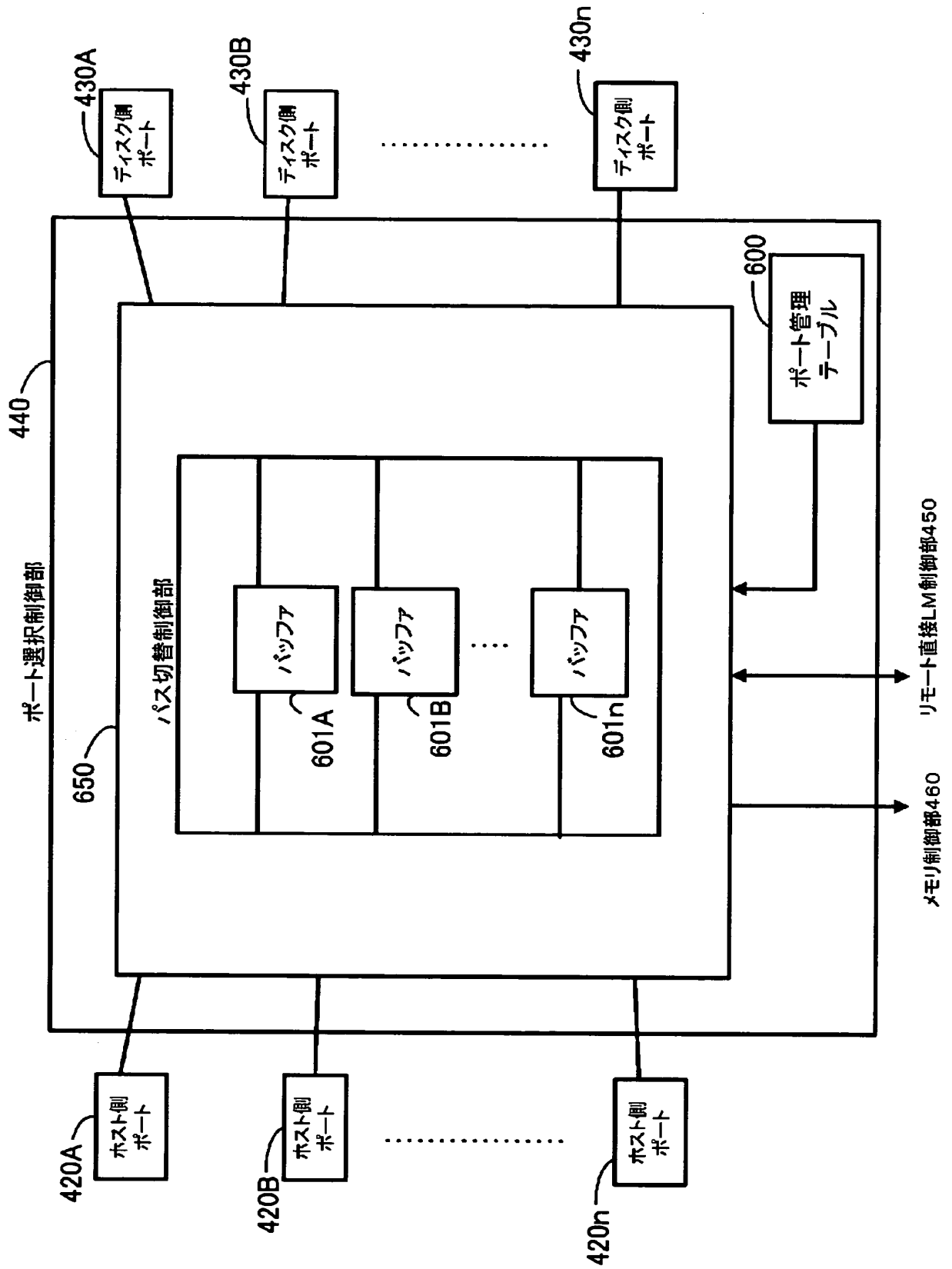
【図 7】

宛先MP 識別情報	バッファ フル	バッファ閾値 オーバー	バッファ エンプティ	バスエラー	ポートコントローラ 内部エラー	シーケンスNo.	CRC
--------------	------------	----------------	---------------	-------	--------------------	----------	-----

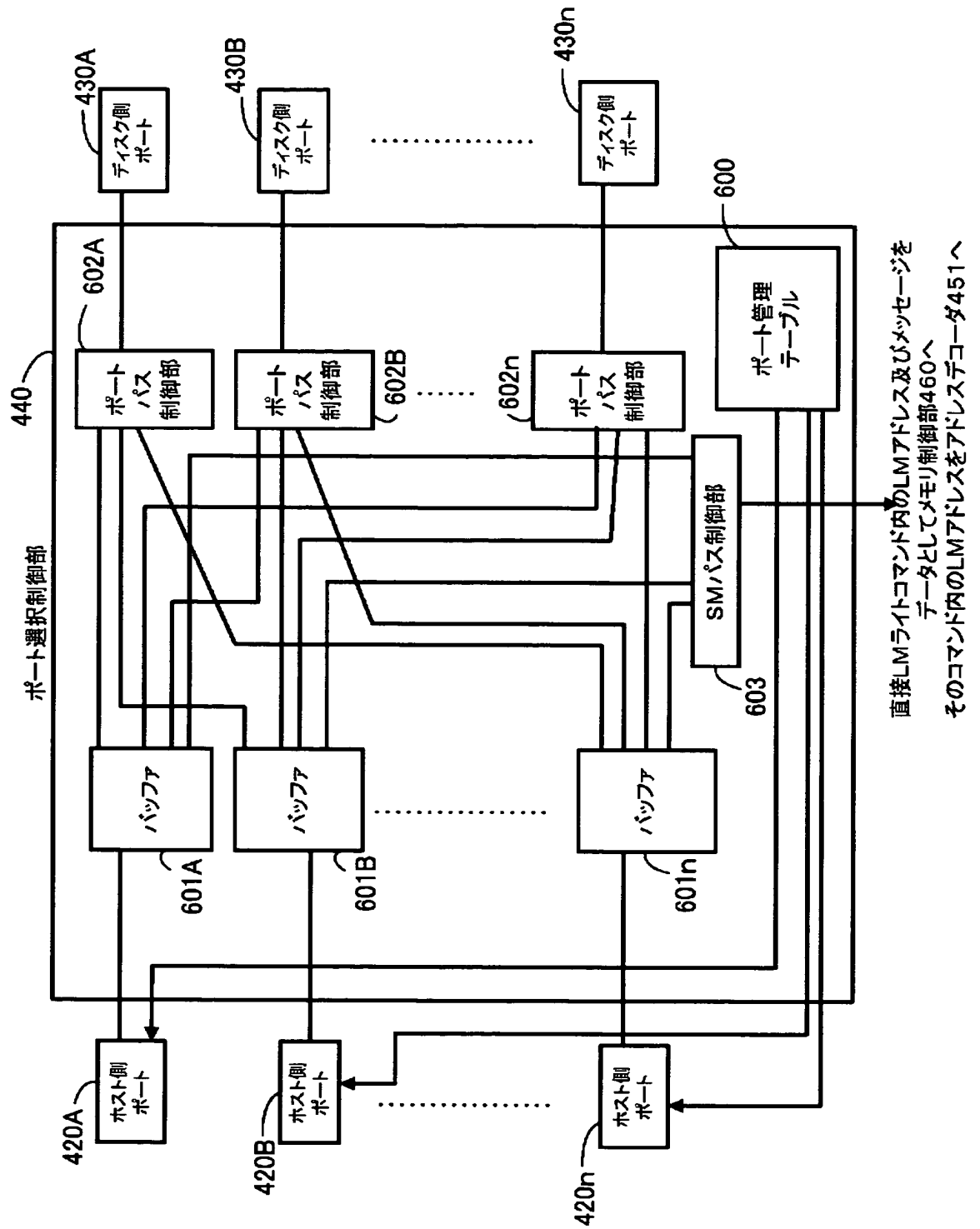
【図 8】

アドレス情報	コマンド種別	ライトデータ	CRC
--------	--------	--------	-----

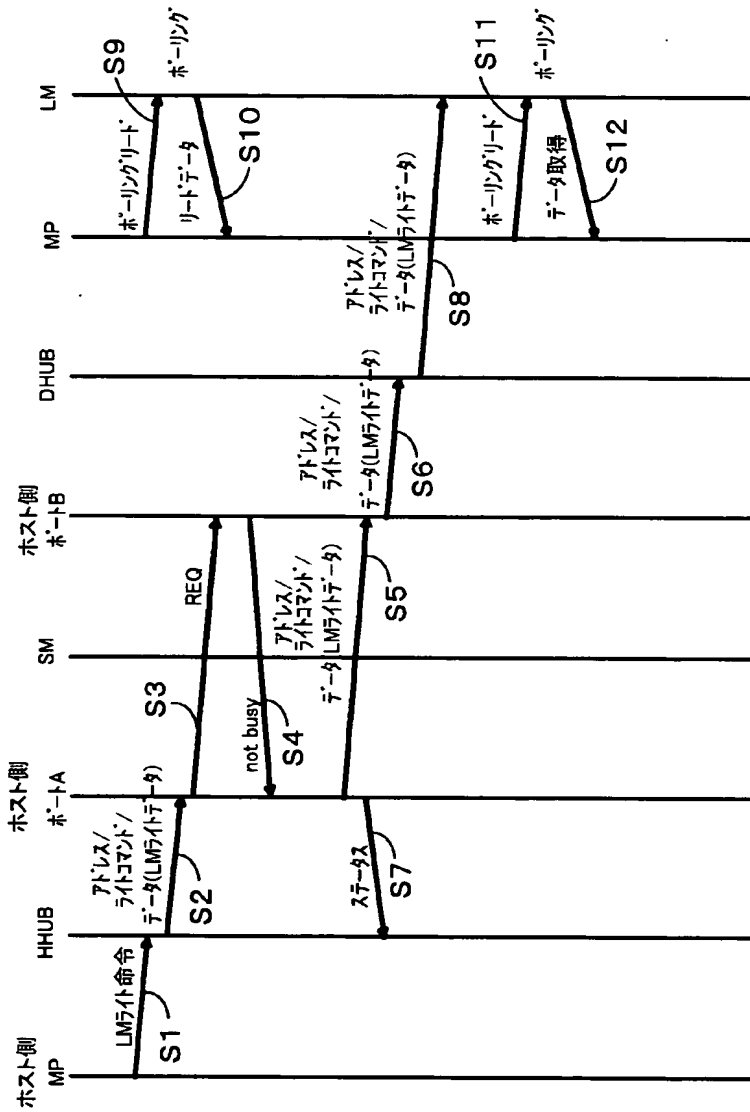
【図 9】



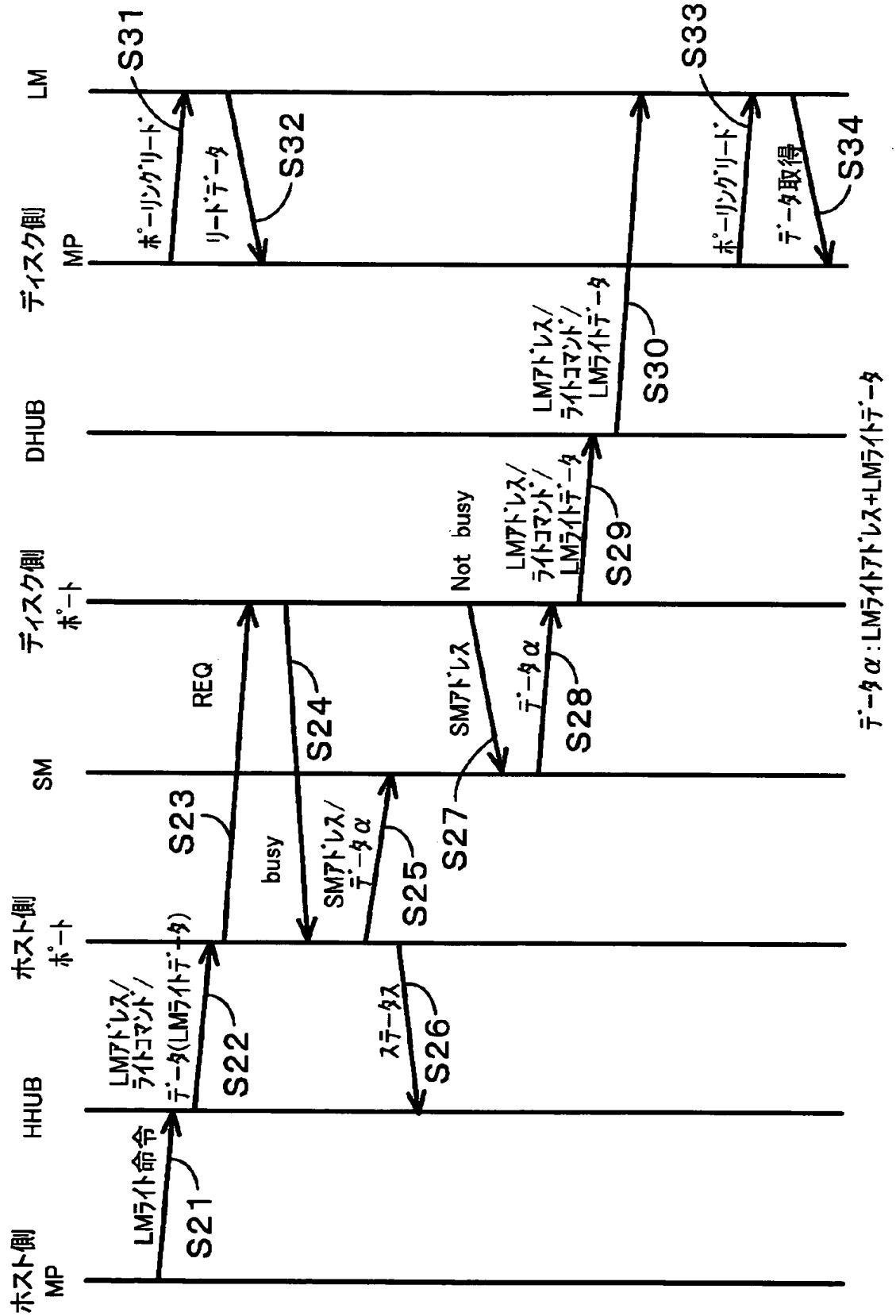
【図 10】



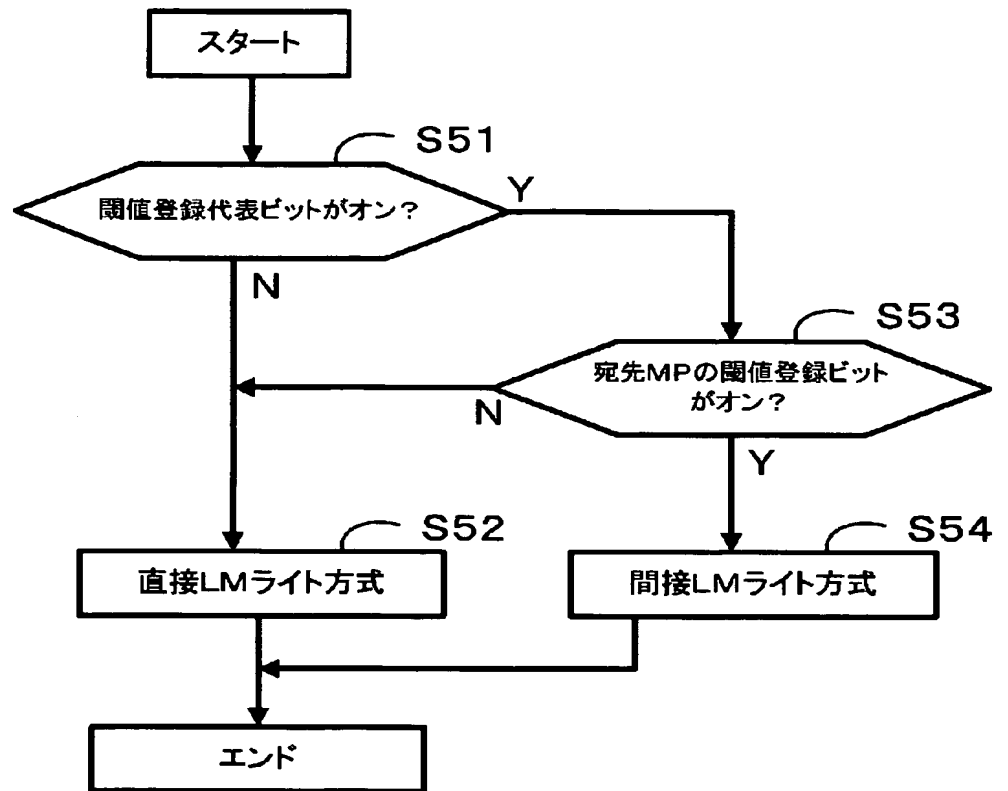
【図 11】



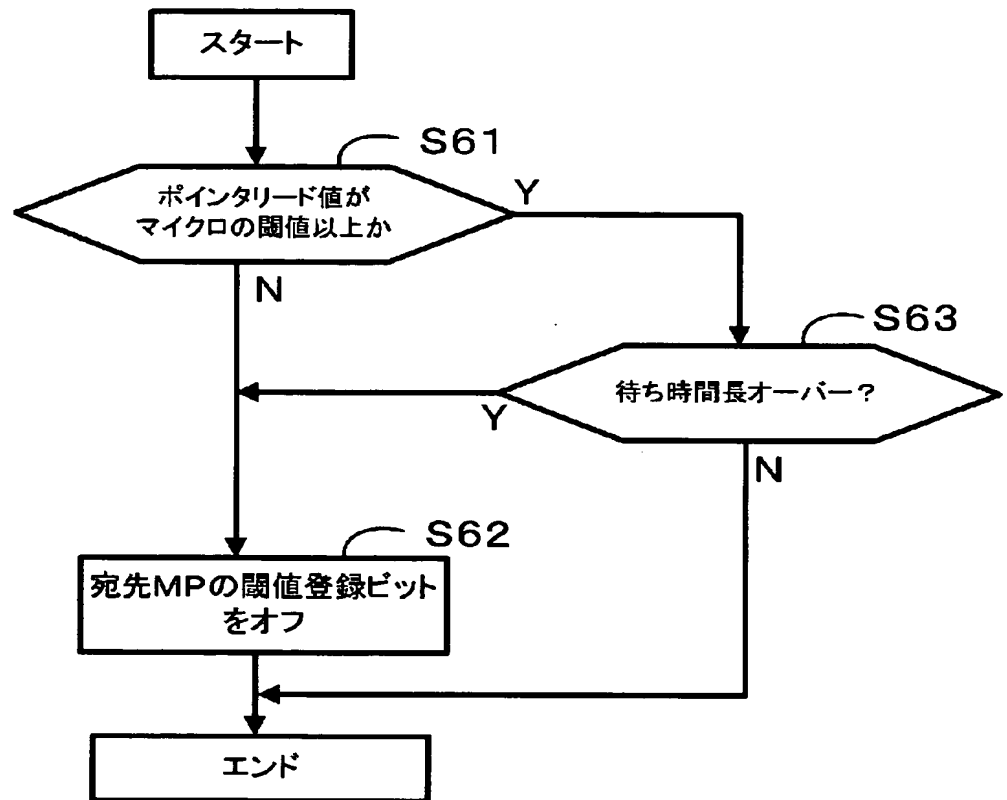
【図 12】



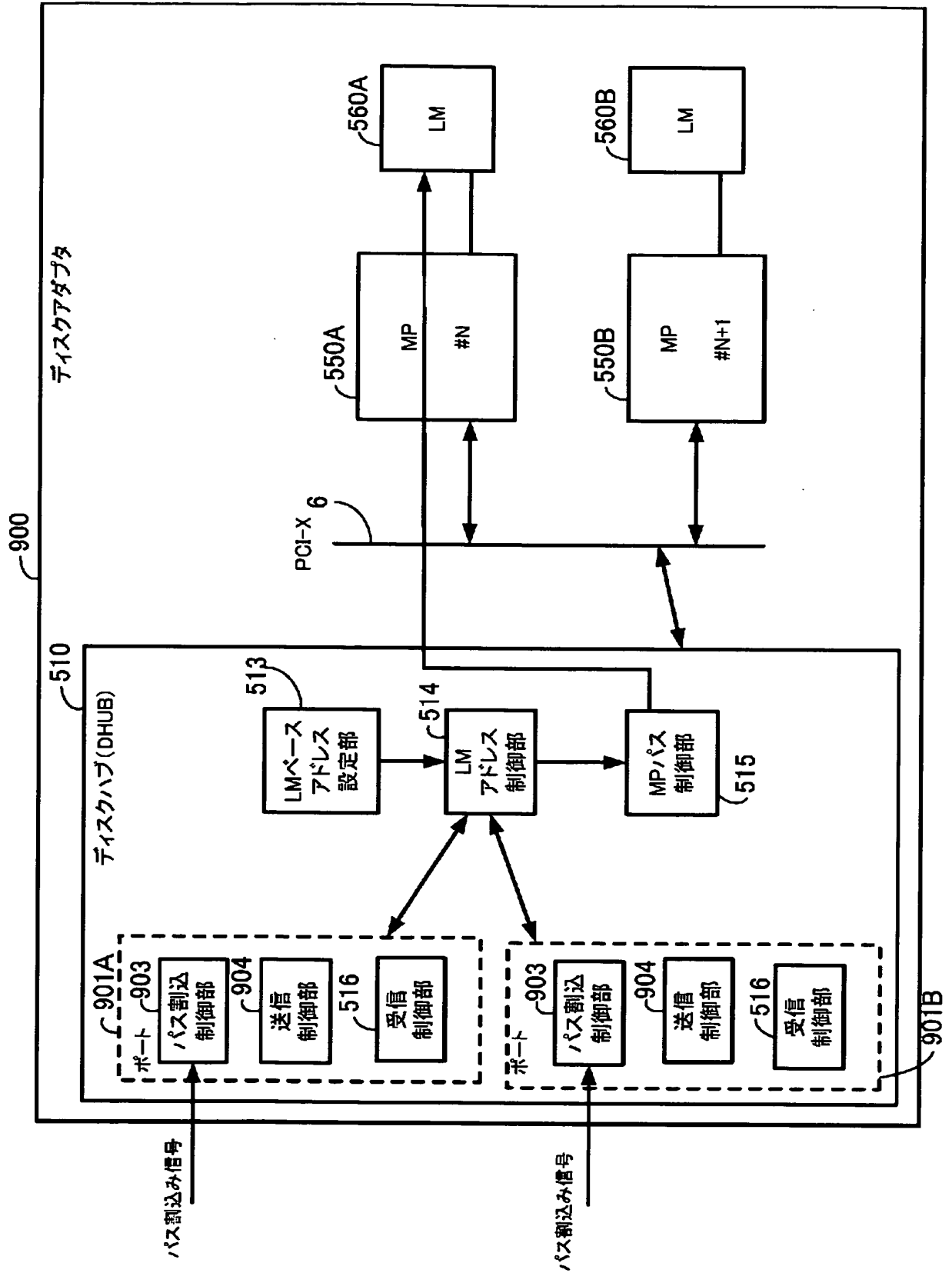
【図13】



【図 14】



【図 15】



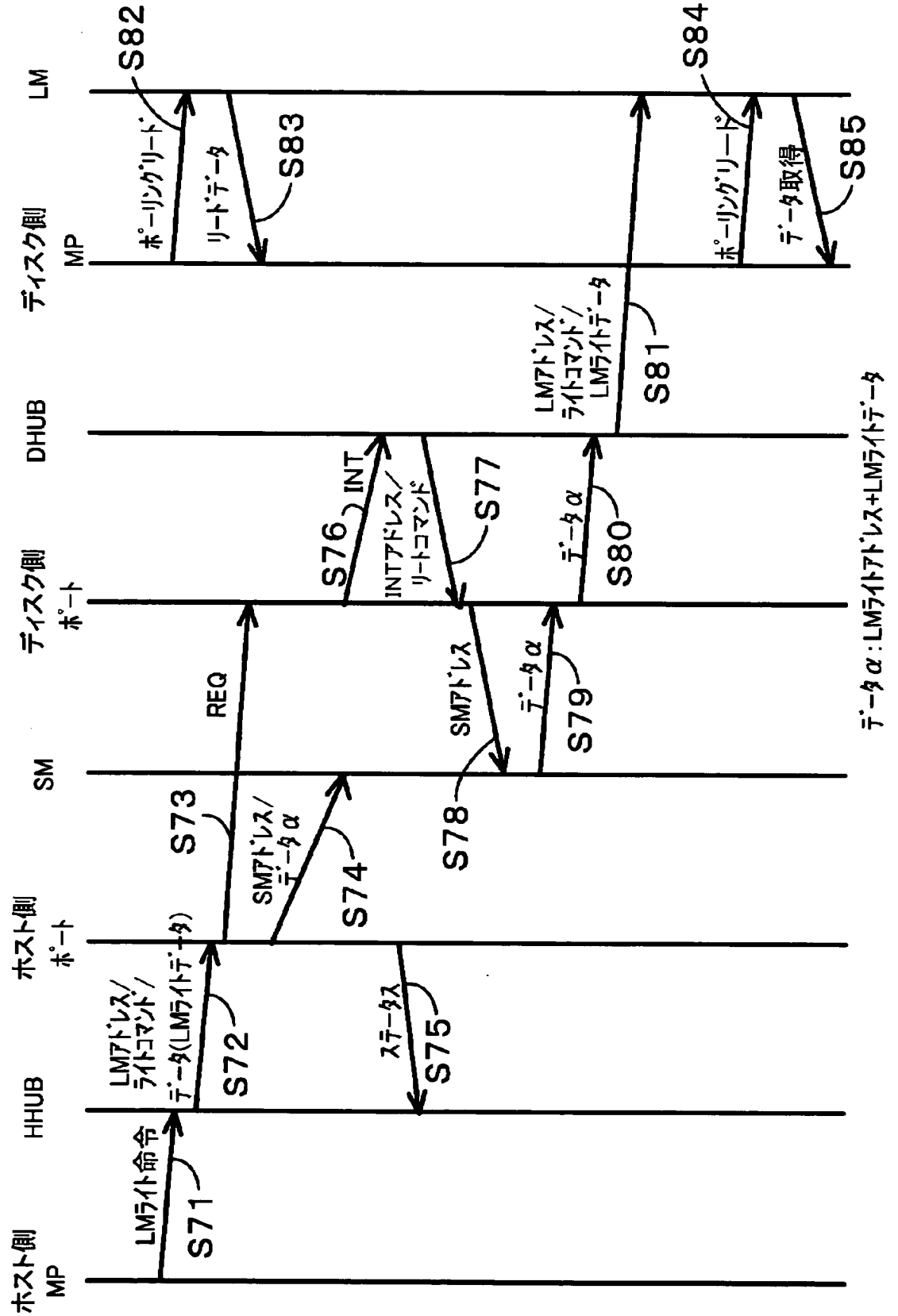
【図16】

アドレス情報 (INTアドレス)	コマンド種別 (リード要求)	CRC
---------------------	-------------------	-----

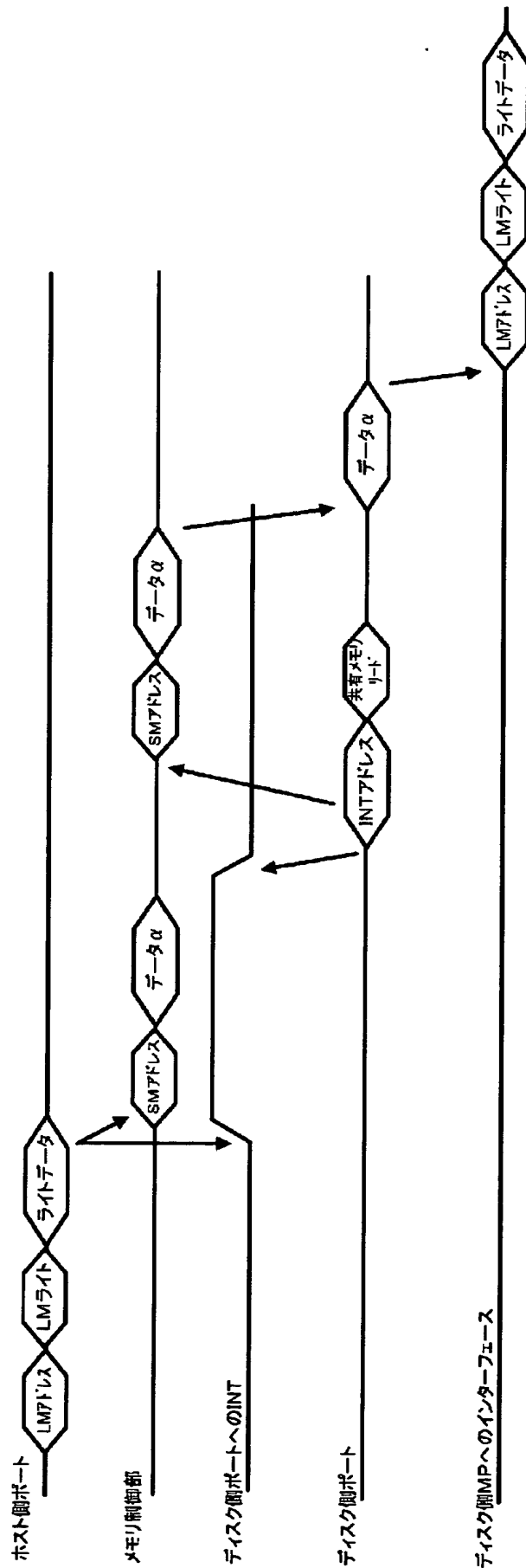
【図17】

Read data (LMアドレス)	Read data (ライトデータ)	CRC
-----------------------	-----------------------	-----

【図18】



【図19】



【図 2 0】

スタート	アダプタ名	他MP #	他LMライトアドレス
0x4000 0000	CHA0	00	0x0000 0000
0x4100 0000		01	0x0000 0000
0x4200 0000	CHA1	02	0x0000 0000
0x4300 0000		03	0x0000 0000
0x4400 0000	CHA2	04	0x0000 0000
0x4500 0000		05	0x0000 0000
0x4600 0000	CHA3	06	0x0000 0000
0x4700 0000		07	0x0000 0000
0x4800 0000	DKA0	08	0x0000 0000
0x4900 0000		09	0x0000 0000
0x4A00 0000	DKA1	0a	0x0000 0000
0x4B00 0000		0b	0x0000 0000
0x4C00 0000	DKA2	0c	0x0000 0000
0x4D00 0000		0d	0x0000 0000
0x4E00 0000	DKA3	0e	0x0000 0000
0x4F00 0000		0f	0x0000 0000

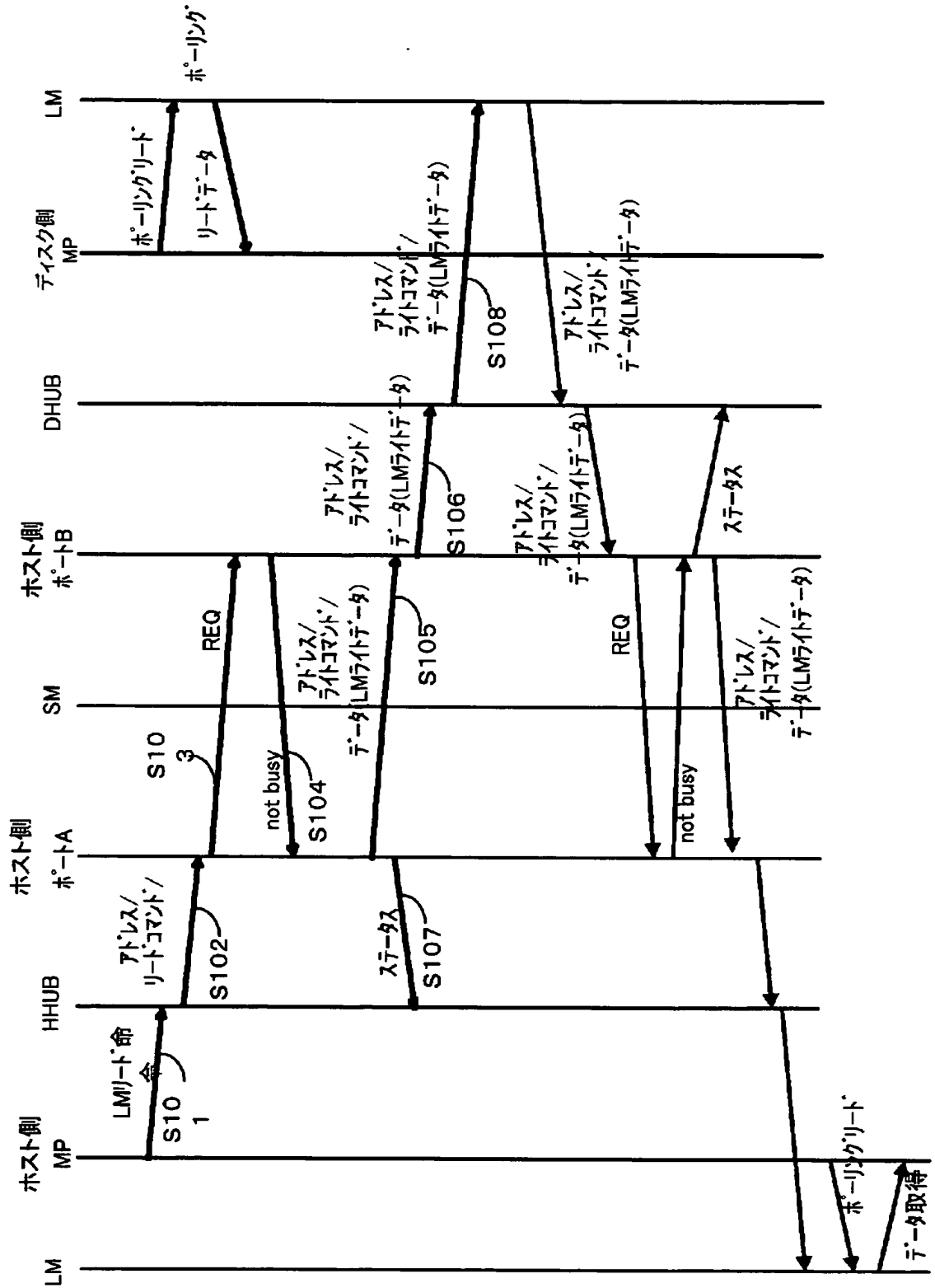
12

MP360 (MP #「00」)

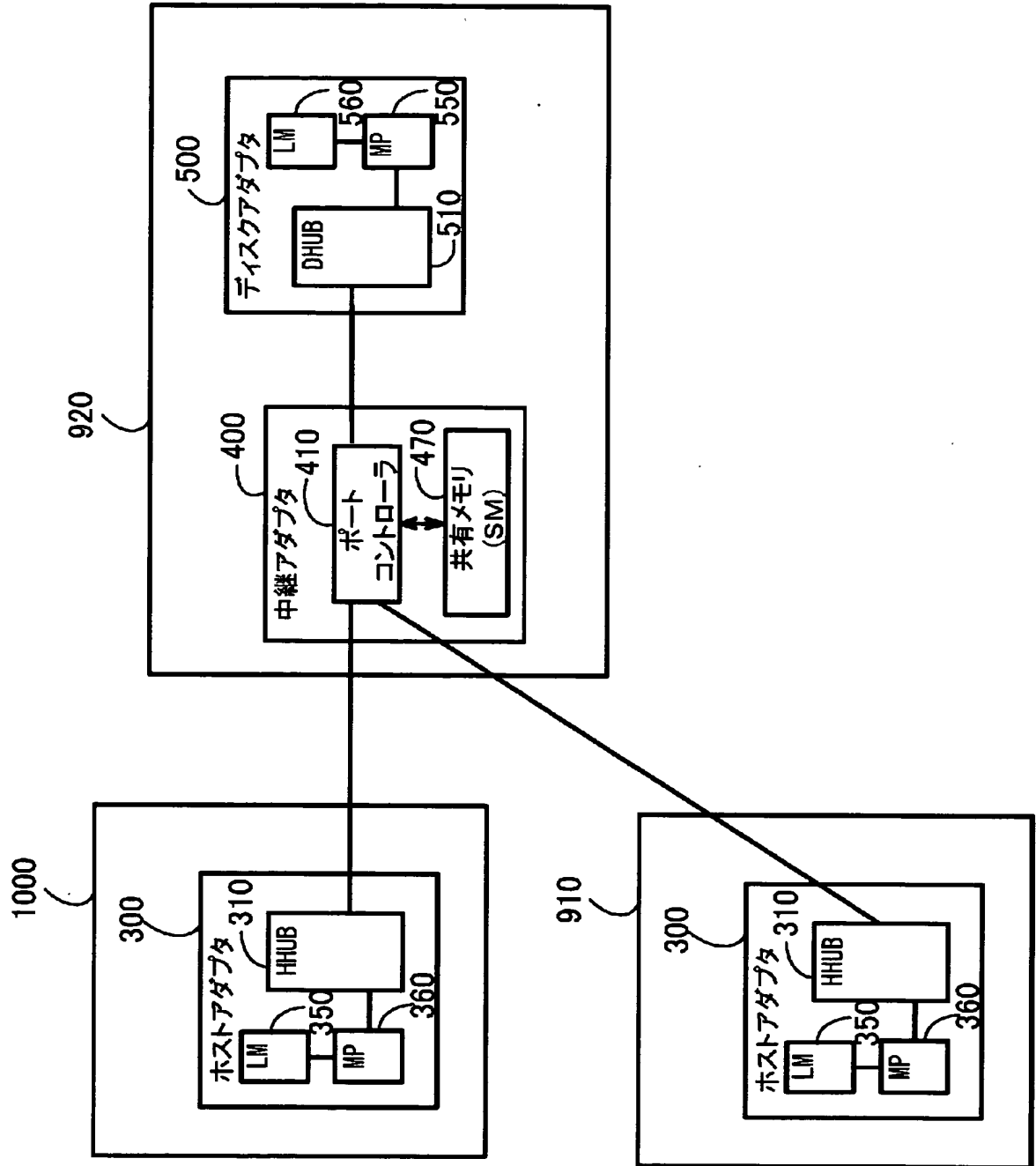
アドレス	用途
0x0000 0000	PCIX 他デバイス
0x1000 0000	SMA I/Oレジスタ
0x2000 0000	送信時SMアドレス情報
0x3000 0000	受信時SMアドレス情報
0x4000 0000	他MPのLMライトアドレス情報
0x5000 0000	自MPのLMライトノードアドレス情報
0x6000 0000	他MPのLMリードアドレス情報

スタート	アダプタ名	他MP #	他LMリードアドレス
0x4000 0000	CHA0	00	0x0000 0000
0x4100 0000		01	0x0000 0000
0x4200 0000	CHA1	02	0x0000 0000
0x4300 0000		03	0x0000 0000
0x4400 0000	CHA2	04	0x0000 0000
0x4500 0000		05	0x0000 0000
0x4600 0000	CHA3	06	0x0000 0000
0x4700 0000		07	0x0000 0000
0x4800 0000	DKA0	08	0x0000 0000
0x4900 0000		09	0x0000 0000
0x4A00 0000	DKA1	0a	0x0000 0000
0x4B00 0000		0b	0x0000 0000
0x4C00 0000	DKA2	0c	0x0000 0000
0x4D00 0000		0d	0x0000 0000
0x4E00 0000	DKA3	0e	0x0000 0000
0x4F00 0000		0f	0x0000 0000

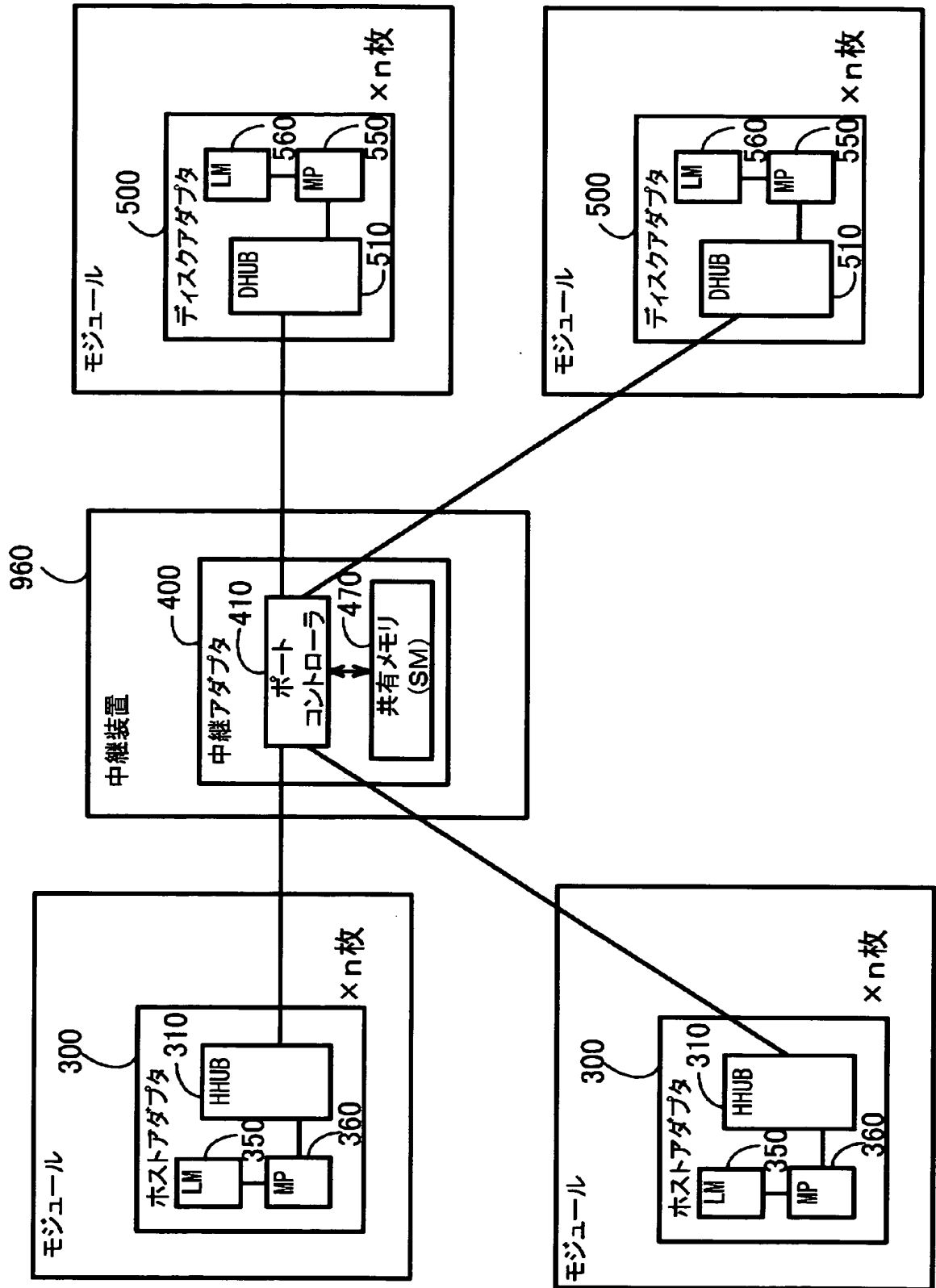
【図21】



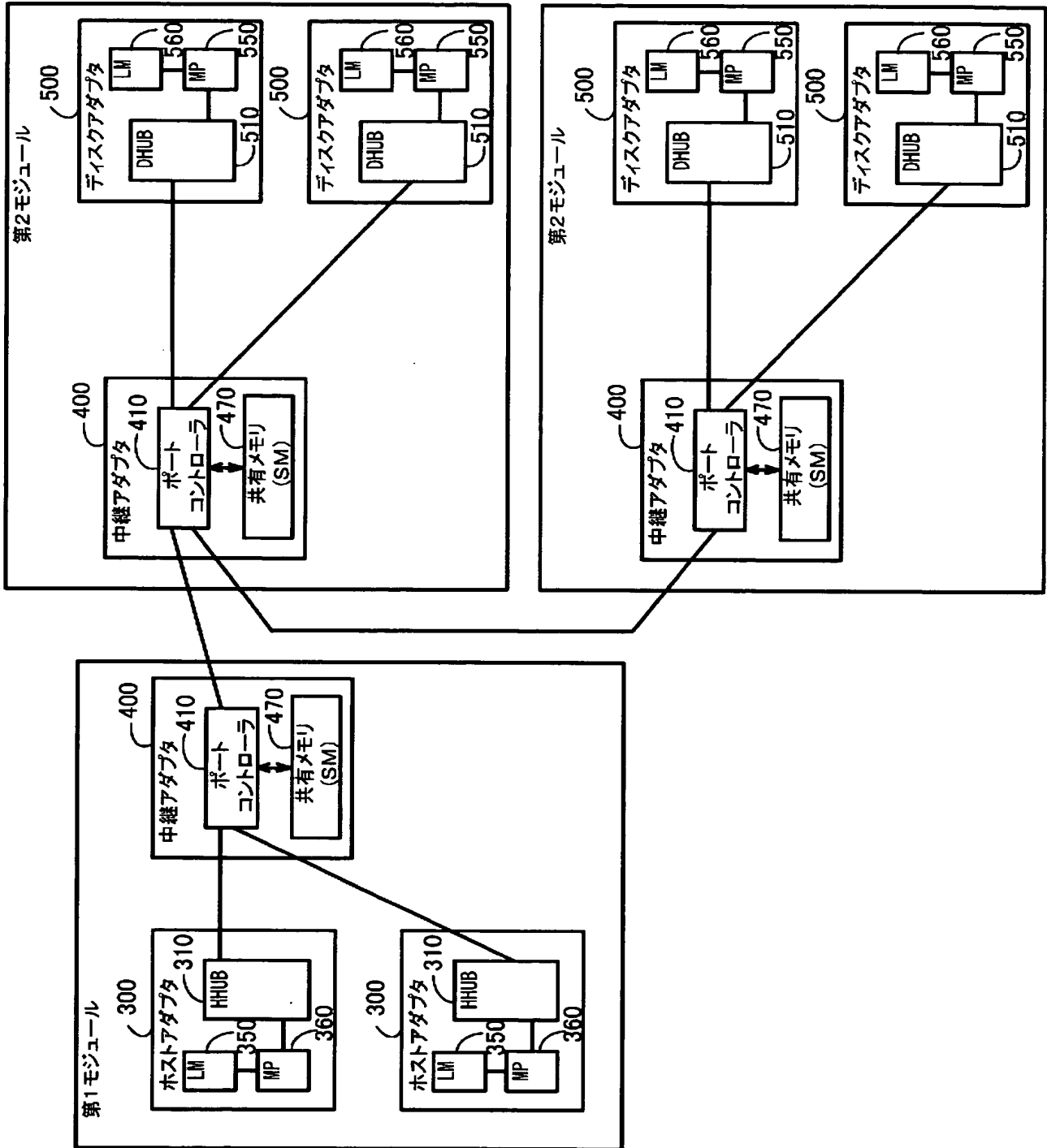
【図23】



【図24】



【図25】



【書類名】 要約書

【要約】

【課題】 受信側プロセッサにとって負担が少ない方法で、プロセッサ間での情報の授受が行えるようにする。

【解決手段】 情報処理装置（100）は、複数のプロセッサを用いて情報を処理する装置であって、1又は複数のローカルメモリ（560）を有する1以上の第1プロセッサ（550）と、ターゲットの第1プロセッサ（550）が有するローカルメモリ（560）にライト情報を直接ライトする1以上の第2プロセッサ（360）とを備える。第2プロセッサ（360）は、第1プロセッサ（550）についてのローカルメモリアドレスが記録されたアドレスマップを記憶しており、そのアドレスマップからローカルメモリアドレスを取得し、その取得したローカルメモリアドレスにライト情報をライトする。

【選択図】 図1

認定・付加情報

特許出願の番号	特願2003-396786
受付番号	50301953985
書類名	特許願
担当官	第七担当上席 0096
作成日	平成15年11月28日

<認定情報・付加情報>

【提出日】 平成15年11月27日

特願 2003-396786

出願人履歴情報

識別番号

[000005108]

1. 変更年月日

1990年 8月31日

[変更理由]

新規登録

住所

東京都千代田区神田駿河台4丁目6番地

氏名

株式会社日立製作所